

Chaotic transport in the homoclinic and heteroclinic tangle regions of quasiperiodically forced two-dimensional dynamical systems

Darin Beigie†, Anthony Leonard‡ and Stephen Wiggins§

† Department of Physics, California Institute of Technology, Pasadena, CA 91125, USA

‡ Department of Aeronautics, California Institute of Technology, Pasadena, CA 91125, USA

§ Department of Applied Mechanics, California Institute of Technology, Pasadena, CA 91125, USA

Received 25 September 1989, in final form 27 November 1990

Accepted by R S MacKay

Abstract. We generalize notions of transport in phase space associated with the classical Poincaré map reduction of a periodically forced two-dimensional system to apply to a sequence of non-autonomous maps derived from a quasiperiodically forced two-dimensional system. We obtain a global picture of the dynamics in homoclinic and heteroclinic tangles using a sequence of time-dependent two-dimensional lobe structures derived from the invariant global stable and unstable manifolds of one or more normally hyperbolic invariant sets in a Poincaré section of an associated autonomous system phase space. The invariant manifold geometry is studied via a generalized Melnikov function (Wiggins). Transport in phase space is specified in terms of two-dimensional lobes mapping from one to another within the sequence of lobe structures, which provides the framework for studying several features of the dynamics associated with chaotic tangles. Instantaneous and average flux with respect to a sequence of cores are quantified analytically in the near-integrable case through the use of the generalized Melnikov function, and the computational prescription for determining flux in systems that are not near-integrable is provided. Other basic transport issues (both in the near-integrable and general case) are studied by a consideration of the geometry of intersections of lobes with each other, with the core, and with level sets of the unperturbed Hamiltonian. We compare transport rates in the single and multiple frequency cases. Though the comparisons depend on choices of normalization, for some reasonable normalizations the average flux of a class of multiple frequency systems is found to be maximal in a certain single frequency limit. The variation of lobe areas in multiple frequency systems, however, gives one the freedom to enhance or diminish aspects of transport over a finite time scale for a fixed average flux. Numerical simulations of lobe structures are presented, using a *double phase slice method* (Beigie *et al*), which provides the basis for *exact* computation of lobe areas and other transport quantities. The chaotic nature of the dynamics is understood in the framework of a *travelling horseshoe map sequence*. The generalized Melnikov function provides a tool to search through parameter space to determine when a quasiperiodically forced system is chaotic. The extension of the analysis to more general time dependences is explained.

AMS classification scheme numbers: 58F27, 58F30, 34A34, 34C35

1. Introduction

Some dynamical systems that exhibit chaotic behaviour evolve according to sufficiently simple vector fields such that one can obtain a fundamental and detailed global picture of the chaotic dynamics. A canonical example is a periodically forced two-dimensional system which can be reduced by time-periodic sampling of trajectories to a two-dimensional Poincaré map (for a detailed account see Guckenheimer and Holmes (1983)). Consider the often studied case where the Poincaré map of the unforced system contains a hyperbolic fixed point connected to itself by a homoclinic orbit, or a pair of hyperbolic fixed points connected to each other by heteroclinic orbits (see figure 1(i)). Periodic forcing drastically alters the dynamics near the unperturbed homoclinic/heteroclinic orbits (also called separatrices). For small enough perturbations the Poincaré map's fixed points persist, and, if their global stable and unstable manifolds intersect once, they will intersect infinitely many times to produce the boundary of a complicated two-dimensional lobe structure (the so called homoclinic/heteroclinic tangle) that is invariant under the Poincaré map (see figure 1(ii)). Lobes map from one to another within the invariant lobe structure with each application of the Poincaré map, and this provides a global picture of the dynamics, i.e. of *transport in phase space*. A variety of physical phenomena can be studied in this context, notable recent examples being advection in fluids (see Rom-Kedar *et al* 1990) and multiphoton ionization (see MacKay and Meiss 1988). However, for periodically forced systems the vector field that evolves the system has a quite simple time dependence, and one is limited to the study of physical systems that can be described by maps. One would like to extend the results to vector fields with as complicated time dependences as possible, while still retaining a fairly complete global picture of the dynamics; this will entail extending the study of maps to the study of *sequences of maps*. Our goal is to generalize some notions of transport in phase space

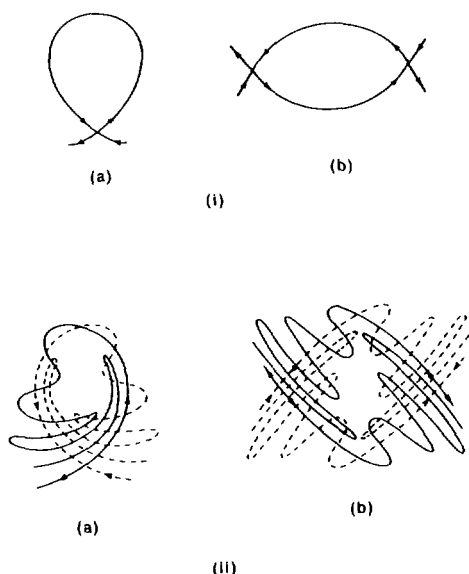


Figure 1. (i) (a) An unperturbed homoclinic orbit, and (b) a pair of unperturbed heteroclinic orbits. (ii) The tangle region of the perturbed system for (a) the homoclinic case, and (b) the heteroclinic case. The dashed and solid lines are stable and unstable manifolds, respectively, of the surviving hyperbolic fixed points.

associated with iterates of a Poincaré map derived from a time-periodic two-dimensional vector field to apply to a bi-infinite sequence of non-autonomous maps derived from a two-dimensional quasiperiodic vector field (i.e. a vector field whose time dependence involves l frequencies, $l \geq 2$). As we shall see, the concepts involved in the generalization are robust and apply to vector fields with more general time dependences. We thus wish to introduce these concepts in the context of a familiar and appealing class of systems.

A wealth of features of the dynamics associated with homoclinic/heteroclinic tangles can be studied in terms of lobes mapping within a set of invariant manifolds, and we introduce two of these features in the time-periodic case before explaining the quasiperiodic extension:

(i) Time-periodic forcing entails the destruction of barriers to transport in phase space, resulting in a net flux across relevant boundaries in phase space and transport from one region to another. The issue of transport in phase space, under iterates of a map, across boundaries formed by segments of stable and unstable manifolds of invariant hyperbolic sets, has been studied by various investigators (see MacKay *et al* 1984, Bensimon and Kadanoff 1984). Rom-Kedar and Wiggins (1990) provide a detailed study of transport under iterates of a two-dimensional map, one which considers the global geometry of the invariant lobe structures that govern the transport, and Wiggins (1990) generalizes this study to a class of k -degree-of-freedom Hamiltonian systems ($2 \leq k < \infty$) with hyperbolicity in two dimensions.

(ii) Within the tangles is a process of repeated stretching and folding to give exponential stretching and sensitive dependence on initial conditions on a non-trivial subset of phase space. The horseshoe map paradigm provides the simplest framework for understanding this phenomenon and the study of lobe behaviour allows one to obtain a global understanding of the stretch properties in chaotic tangles (see Beigie *et al* 1991a).

The addition of even one more frequency to a vector field's time dependence requires a fundamental departure from the time-periodic analysis: the evolution of trajectories from one time sample to the next (for whatever choice of sampling) must now in general be described by a different map for each time sample. One can thus no longer construct an invariant two-dimensional lobe structure upon which to study transport, or invoke the standard horseshoe map construction to establish that there is chaotic motion. How then does one obtain a detailed global picture of the dynamics? To study the dynamics of a non-autonomous system that evolves according to a two-dimensional, l frequency vector field, we consider in section 2 an associated $(l+2)$ -dimensional autonomous system in order to construct an invariant $(l+1)$ -dimensional lobe structure in a Poincaré section of the enlarged phase space (the lobe boundaries being formed out of intersecting global stable and unstable manifolds of normally hyperbolic invariant $(l-1)$ -tori in the defined Poincaré section). We derive from this invariant lobe structure a sequence of two-dimensional time-dependent lobe structures in the non-autonomous system phase space that is used to analyse transport under a sequence of non-autonomous maps. For concreteness, we focus on quasiperiodically forced systems whose unforced solution contains in the Poincaré section a single normally hyperbolic invariant $(l-1)$ -torus connected to itself by a homoclinic orbit, or a pair of such tori connected by two or more heteroclinic orbits. The appendix defines an example system, an oscillating vortex pair (ovp) flow that will be used to illustrate our discussion (see Rom-Kedar *et al* (1990) for the time-periodic case). A generalized Melnikov function (Wiggins

1987, 1988) will repeatedly help us understand the nature of the invariant $(l+1)$ -dimensional lobe structure, which has more complicated geometrical considerations than in the single frequency case. In particular, in section 2 we study the zero sets of the Melnikov function to determine the geometry of intersections of stable and unstable manifolds, and hence the geometry of the invariant lobes. Section 3 describes how the lobes of one time sample of the non-autonomous system map to the lobes of the next time sample; we define a sequence of time-dependent cores and turnstile lobes, explain the procedure for determining the number of turnstile lobes for each time sample, and illustrate this procedure with some two frequency examples. The turnstile lobes are the mechanism for entrainment and detrainment relative to the core sequence (since for simplicity we are dealing with a single core sequence, we borrow the fluid mechanical terminology of entrainment and detrainment to make the discussion of transport in and out of the core sequence, respectively, more apparent). The main extension from the previously mentioned studies of transport across boundaries is the notion of a sequence of time-dependent cores, and we discuss how such a notion is reasonable, and how the stable and unstable manifolds in the Poincaré section of the autonomous system phase space provide the natural definition of a sequence of time-dependent cores. We conclude section 3 by presenting a computer simulation of a sequence of lobe structures and of lobes mapping from one to another within this sequence, and discussing qualitatively the essential features of transport in the time-dependent tangle region. The sampling method used for the simulations, which we refer to as a *double phase slice method* (Beigie et al 1991b), contrasts with a previous suggestion by Moon and Holmes (1985) for sampling quasiperiodically forced systems to find underlying structure and provides the basis for *exact* numerical computation of lobe areas and other transport quantities. The next two sections examine more closely the features discussed qualitatively in section 3. Section 4 focuses on entrainment and detrainment relative to a sequence of two-dimensional time-dependent cores in the non-autonomous system phase space. We calculate analytically instantaneous and average flux with respect to the sequence of cores in near-integrable systems, and present the computational prescription for calculating these quantities when the system is not near-integrable. We compare average flux in the single and multiple frequency cases; though the comparisons depend on the normalization of perturbation amplitudes we shall see how, for reasonable normalizations, the average flux for a class of multiple frequency systems is maximal in a particular single frequency limit. We explain how answering basic questions about transport in phase space (both in the near-integrable and general case) necessitates a consideration of more than just flux, and then study such issues as the geometry of intersections of lobes with each other, with the core, and with level sets of the unperturbed Hamiltonian. For example, we will see that even though average flux may decrease in quasiperiodic systems, other factors can outweigh this effect to make it easier for inner level sets of the unperturbed Hamiltonian to escape the core. More generally, *the variation of lobe areas gives one the freedom to enhance or diminish transport properties over finite times for a fixed average flux*. Section 5 deals with the description of chaos associated with the stretching and folding of lobes in the framework of a *travelling horseshoe map sequence*. The chaotic dynamics of the non-autonomous system is understood relative to a time-dependent Cantor set of points. Other studies, with different methods, of the chaotic nature of the dynamics under perturbations with time dependences more complicated than periodic, or,

equivalently, under a sequence of non-autonomous maps, are given by Scheurle (1986), Stoffer (1988a, b), and Meyer and Sell (1989). The generalized Melnikov function allows one to search in the parameter space of the quasiperiodic system to determine when one can and cannot have chaos (see Ide and Wiggins 1989). Such a search can have simple, practical consequences: for example, we shall see for our systems how a relative phase shift in the forcing can lead to suppression of chaos in a class of multiple frequency vector fields. We conclude in section 6 with an explanation of how the analysis extends to vector fields with more general time dependences than quasiperiodic. We will emphasize throughout the paper the two frequency case as a workable and illuminating example of transport in phase space under a bi-infinite sequence of non-repeating maps.

2. Defining the lobe structures and studying their geometry

We consider two-dimensional non-autonomous systems of the form

$$\dot{x} = JDH(x) + \varepsilon g(x, t; \mu, \varepsilon) \quad (2.1)$$

where $x \in \mathbb{R}^2$, $H: X \mapsto \mathbb{R}^1$ is $C^{r+1}(r \geq 2)$ on some open set $X \subset \mathbb{R}^2$, $g: X \times \mathbb{R}^1 \times V \times \mathbb{R}^1 \mapsto \mathbb{R}^2$ is $C^r(r \geq 2)$ with parameter space $V \subset \mathbb{R}^p$ an open set, $\varepsilon \in \mathbb{R}^1$, and

$$J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \quad DH(x) = \begin{pmatrix} \partial H / \partial x_1 \\ \partial H / \partial x_2 \end{pmatrix}.$$

Equation (2.1) represents a class of perturbed single degree-of-freedom Hamiltonian systems (note that the perturbation $\varepsilon g(x, t; \mu, \varepsilon)$ may or may not be Hamiltonian). We restrict our interest to perturbations periodic or quasiperiodic in time for fixed $x \in X$, $\mu \in V$, $\varepsilon \in \mathbb{R}^1$, and hence we can rewrite the governing equation as

$$\dot{x} = JDH(x) + \varepsilon g(x, \theta_1(t), \theta_2(t), \dots, \theta_l(t); \mu, \varepsilon) \quad (2.2)$$

where $\theta_i(t) = \omega_i t + \theta_{i0}$ for $i = 1$ to l (some positive integer ≥ 1), and g is 2π -periodic in each of $\theta_i(t)$. Though we focus on quasiperiodic systems ($l \geq 2$), it should be clear as we proceed how our $l \geq 2$ analysis applies to the familiar time-periodic ($l = 1$) case, and we will refer from time to time to the single frequency case.

Let $x(t) = \phi(t, t_0, x_0)$ denote the solution to (2.2) which satisfies the initial condition $\phi(t_0, t_0, x_0) = x_0$. To study the dynamics of (2.2), it will be advantageous to sample the trajectories at discrete time intervals, the interval being one of the perturbation periods, say $2\pi/\omega_l$. For time-periodic vector fields, it is well understood that periodic sampling filters out redundant dynamical information to obtain a simpler underlying structure with which to study the dynamics; for quasiperiodic vector fields, periodic sampling will in a similar way simplify the underlying structure with which the dynamics is studied. The evolution of system (2.2) from time $t = (2\pi/\omega_l)n$ ($n \in \mathbb{Z}$) to time $t = (2\pi/\omega_l)(n+1)$ defines a map on X :

$$T_\varepsilon(\cdot; n): X \mapsto X$$

$$x \mapsto \phi\left(\frac{2\pi}{\omega_l}(n+1), \frac{2\pi}{\omega_l}n, x\right).$$

Our goal then is to study the dynamics in X under the bi-infinite sequence of non-autonomous maps $\{T_\varepsilon(\cdot; n); n \in \mathbb{Z}\}$. To do this, we employ a standard technique for the study of non-autonomous systems (see Guckenheimer and Holmes 1983 or Wiggins 1988), that is to examine an associated autonomous system in order to construct invariant manifolds that are fixed in the enlarged phase space. These manifolds form barriers which constrain the motion in the autonomous system phase space and hence provide a global picture of the dynamics. Having obtained these manifolds, we will use them to define for each time $t = (2\pi/\omega_l)n$ a set of curves in X that will be used in later sections to study the dynamics in X under the map sequence $\{T_\varepsilon(\cdot; n); n \in \mathbb{Z}\}$.

2.1. Constructing an invariant lobe structure and studying its geometry

We rewrite equation (2.2) in the autonomous form

$$\begin{aligned}\dot{x} &= JDH(x) + \varepsilon g(x, \theta_1, \theta_2, \dots, \theta_l; \mu, \varepsilon) \\ \dot{\theta}_i &= \omega_i \quad i = 1, 2, \dots, l.\end{aligned}\tag{2.3}$$

The autonomous system phase space is $\mathbb{R}^2 \times T^l$ (where T^l is the l -torus). The expression

$$(x(t), \theta_1(t), \theta_2(t), \dots, \theta_l(t)) = (\phi(t, t_0=0, x_0), \omega_1 t + \theta_{1_0}, \omega_2 t + \theta_{2_0}, \dots, \omega_l t + \theta_{l_0})\tag{2.4}$$

solves equation (2.3). Let us denote this solution by $\Phi(t, t_0=0, x_0, \theta_{1_0}, \theta_{2_0}, \dots, \theta_{l_0})$. From equation (2.4), the trajectory of (2.2) through x_0 at time $t=0$ is the x component of the trajectory of (2.3) through $(x_0, \theta_{1_0}, \theta_{2_0}, \dots, \theta_{l_0})$ at time $t=0$, and hence the study of trajectories of the non-autonomous system is related to the study of the appropriate trajectories of the autonomous system.

We define a global cross section of the autonomous system phase space

$$\Sigma^{\theta_{l_0}} = \{(x, \theta_1, \theta_2, \dots, \theta_l) \mid \theta_l = \theta_{l_0}\}$$

and the associated Poincaré map generated by the flow of (2.3) is given by

$$P_\varepsilon : \Sigma^{\theta_{l_0}} \mapsto \Sigma^{\theta_{l_0}}$$

$$\left(x, \theta_0 + 2\pi \frac{\omega}{\omega_l} n\right) \mapsto \left(\phi\left(\frac{2\pi}{\omega_l}(n+1), \frac{2\pi}{\omega_l} n, x\right), \theta_0 + 2\pi \frac{\omega}{\omega_l}(n+1)\right)\tag{2.5}$$

where $\theta_0 \equiv (\theta_{1_0}, \theta_{2_0}, \dots, \theta_{(l-1)_0})$ and $\omega \equiv (\omega_1, \omega_2, \dots, \omega_{(l-1)})$. Note that the map is independent of n since it is derived from an autonomous system. Studying the flow of (2.3) via this Poincaré map is equivalent to sampling the trajectories of (2.3) at time intervals equal to $2\pi/\omega_l$.

We study cases in which the unperturbed ($\varepsilon = 0$) system (2.2) has: (a) a hyperbolic fixed point x_h which is connected to itself by a homoclinic trajectory $x_h(t)$, or (b) a pair of hyperbolic fixed points x_h^a, x_h^b connected to each other by heteroclinic trajectories $x_h^i(t)$ (i labels different trajectories). From these basic structures, the results will generalize to an arbitrary number of hyperbolic fixed points, each connected to itself and/or other fixed points by homoclinic/heteroclinic trajectories. The Poincaré map of the unperturbed homoclinic system has a normally hyperbolic

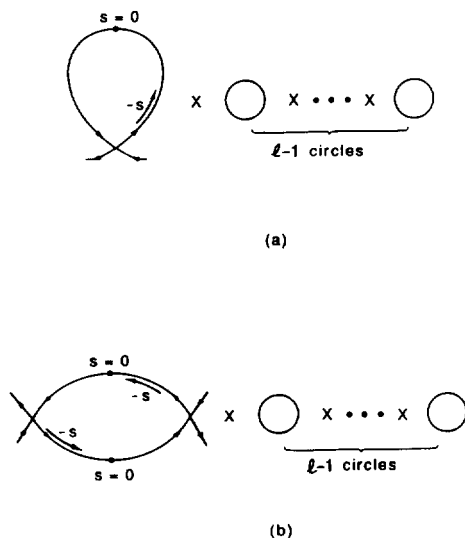


Figure 2. (a) A homoclinic manifold and (b) two heteroclinic manifolds in the Poincaré section Σ^{θ_0} .

invariant $(l-1)$ -torus of the form

$$\tau_0 = \{(x, \theta_1, \theta_2, \dots, \theta_l) \mid x = x_h, \theta_l = \theta_{l_0}\}$$

whose l -dimensional stable and unstable manifolds, denoted $W^s(\tau_0)$ and $W^u(\tau_0)$ respectively, coincide along the l -dimensional homoclinic manifold given by

$$W^s(\tau_0) \cap W^u(\tau_0) = \{(x, \theta_1, \theta_2, \dots, \theta_l) \mid x = x_h(t), \theta_l = \theta_{l_0}, t \in \mathbb{R}\}.$$

The heteroclinic case is understood similarly. Note that ‘normal hyperbolicity’ refers to the fact that expansion and contraction rates normal to τ_0 dominate those tangent to it, as to be expected since the θ_i motion is only linear in time. Figure 2 shows the manifolds for the l frequency case (in the forthcoming diagrams, we will not keep presenting both the homoclinic and heteroclinic case, but rather show one or the other interchangeably). The parameter s specifies the location in X along the invariant unperturbed homoclinic or heteroclinic manifold: it represents the time it takes for the point $x(-s)$ on the unperturbed manifold in X to move according to the flow (2.3) to the point $x(s=0)$ (defined to be the spatial midpoint of the unperturbed manifold in X).

To study the perturbed system we invoke the following persistence result for small perturbations.

Proposition 2.1. If $P_{\varepsilon=0}$ contains a normally hyperbolic invariant $(l-1)$ -torus τ_0 , then for ε sufficiently small, P_ε possesses a C^r normally hyperbolic invariant $(l-1)$ -torus, τ_ε , whose local l -dimensional C^r stable and unstable manifolds, $W_{\text{loc}}^s(\tau_\varepsilon)$ and $W_{\text{loc}}^u(\tau_\varepsilon)$, are C^r ε -close to $W_{\text{loc}}^s(\tau_0)$ and $W_{\text{loc}}^u(\tau_0)$, respectively.

Proof. This is an immediate consequence of the persistence theory for normally hyperbolic invariant manifolds. See Fenichel (1971), Hirsch *et al* (1977), or Wiggins (1988).

In the usual way we can define global stable and unstable manifolds, $W^s(\tau_\varepsilon)$ and $W^u(\tau_\varepsilon)$, from the local ones. The perturbation causes these manifolds to undergo a global bifurcation, and the geometrical possibilities of this bifurcation are far richer than in the time-periodic case. A generalized Melnikov function (Wiggins 1988) provides a tool for studying manifold geometry:

$$M(s, \theta_1, \dots, \theta_{(l-1)}, \theta_{l_0}; \nu) = \int_{-\infty}^{\infty} DH(x_h(t)) \cdot g(x_h(t), \omega_1 t + (\omega_1 s + \theta_1), \dots, \omega_{(l-1)} t + (\omega_{(l-1)} s + \theta_{(l-1)}), \omega_l t + (\omega_l s + \theta_{l_0}); \mu, \varepsilon = 0) dt, \quad (2.6)$$

where $(s, \theta_1, \dots, \theta_{(l-1)})$ specify a point on $W^s(\tau_0^a) \cap W^u(\tau_0^b)$ (and are thus constant in the integral expression), and $\nu = (\mu, \omega_1, \omega_2, \dots, \omega_l)$. Whenever we write τ_0^a and τ_0^b , it is understood that in the homoclinic case $a = b$. Further, note that we leave out the i superscript on $x_h(t)$; in the heteroclinic case, it is understood that $x_h(t)$ represents one orbit out of a set of possible heteroclinic orbits. In the Poincaré section $\Sigma^{\theta_{l_0}}$, the first order in ε approximation to the signed distance between the global stable and unstable manifolds is

$$d(s, \theta_1, \dots, \theta_{(l-1)}; \nu, \varepsilon) = \frac{\varepsilon M(s, \theta_1, \dots, \theta_{(l-1)}, \theta_{l_0}; \nu)}{\|DH(x_h(-s))\|} + O(\varepsilon^2) \quad (2.7)$$

(see Wiggins 1988). More specifically, d is the signed distance, measured along the normal to the unperturbed homoclinic or heteroclinic orbit, between the 'first' intersection of $W^u(\tau_\varepsilon^b)$ with the normal and the 'first' intersection of $W^s(\tau_\varepsilon^a)$ with the normal (see figure 3). 'First' is understood relative to moving along $W^u(\tau_\varepsilon^b)$ normally away from τ_ε^b and along $W^s(\tau_\varepsilon^a)$ normally away from τ_ε^a . Note that $\|DH(x_h(-s))\| \rightarrow 0$ as $s \rightarrow \pm\infty$, so the distance function blows up with $|s|$, making the first order in ε approximation in (2.7) a poor one for sufficiently large $|s|$. The approximation will be valid however over a limited domain: the manifolds $W^s(\tau_\varepsilon^a)$, $W^u(\tau_\varepsilon^b)$ can be approximated uniformly on semi-infinite time intervals (see Guckenheimer and Holmes 1983, Wiggins 1988), and on these time intervals the manifolds move only an $O(\varepsilon)$ amount from the unperturbed manifolds. Hence for sufficiently small perturbations the manifold separations will be well approximated by the $O(\varepsilon)$ term over a finite domain.

For illustration purposes we consider systems with perturbations of the form

$$g(x, t; \mu, \varepsilon) = F_0(x; \mu) + \mathcal{F}(x; \mu) \sum_{i=1}^l F_i(\mu, \omega_i) \sin(\omega_i t + \theta_{i_0}) + O(\varepsilon). \quad (2.8)$$

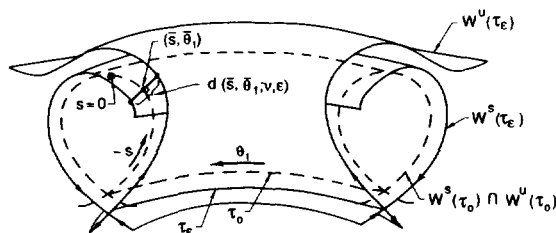


Figure 3. The signed distance function between the stable and unstable manifolds in $\Sigma^{\theta_{l_0}}$ for a perturbed two frequency homoclinic system (refer to Wiggins (1988) for determining the sign of the Melnikov function). A cut-away half view is shown.

The generalized Melnikov function associated with these perturbations has the form

$$\begin{aligned} M(s, \theta_1, \dots, \theta_{l-1}, \theta_{l_0}; \nu) = & A_0(\mu) + \sum_{i=1}^{l-1} A_i(\mu, \omega_i) \sin(\omega_i s + \theta_i) \\ & + A_l(\mu, \omega_l) \sin(\omega_l s + \theta_{l_0}) + \sum_{i=1}^{l-1} B_i(\mu, \omega_i) \cos(\omega_i s + \theta_i) \\ & + B_l(\mu, \omega_l) \cos(\omega_l s + \theta_{l_0}). \end{aligned} \quad (2.9)$$

In particular, the appendix gives the generalized Melnikov function for the ovp flow. We shall refer to the ratio of the Melnikov function amplitude, $A_i(\mu, \omega_i)$ or $B_i(\mu, \omega_i)$, to the corresponding relative perturbation amplitude, $F_i(\mu, \omega_i)$, as a *relative scaling factor*:

$$\frac{A_i(\mu, \omega_i)}{F_i(\mu, \omega_i)} \quad \text{or} \quad \frac{B_i(\mu, \omega_i)}{F_i(\mu, \omega_i)} \equiv \text{relative scaling factor associated with frequency } \omega_i.$$

The relative scaling factors are in general *frequency dependent*, which will be relevant to the study of how manifold separation and transport rates depend on single and multiple frequency forcing. As illustrated in the appendix, all scaling factors are determined by two *relative scaling functions* (one for the A 's, one for the B 's) which are derived from the generalized Melnikov function (in fact all the B 's vanish for the ovp flow, so there is only a single scaling function for this example). We should point out that if the perturbation amplitudes in (2.8) each had a different spatial dependence, the corresponding Melnikov function would still be expressed in the form (2.9), since the spatial dependences are integrated out. We choose for concreteness, however, to refer to the class of perturbations in (2.8) for illustration (whose attractive features include being able to define relative scaling factors).

To define and study an invariant lobe structure necessitates a consideration of the geometry of intersection manifolds of $W^s(\tau_\epsilon^a)$ and $W^u(\tau_\epsilon^b)$, and the following theorem allows one to use the Melnikov function for this consideration.

Theorem 2.2. Suppose there exists a point $(\bar{s}, \bar{\theta}_1, \dots, \bar{\theta}_{(l-1)})$ for some $\bar{\nu}$ such that

- (1) $M(\bar{s}, \bar{\theta}_1, \dots, \bar{\theta}_{(l-1)}, \theta_{l_0}; \bar{\nu}) = 0$,
- (2) $D_{(s, \theta_1, \dots, \theta_{(l-1)})} M(\bar{s}, \bar{\theta}_1, \dots, \bar{\theta}_{(l-1)}, \theta_{l_0}; \bar{\nu})$ is of rank one.

Then for ϵ sufficiently small $W^s(\tau_\epsilon^a)$ and $W^u(\tau_\epsilon^b)$ intersect transversely near (within $O(\epsilon)$) this point, and this intersection can be continued to an $(l-1)$ -dimensional manifold in $\Sigma^{\theta_{l_0}}$.

Proof. This is a straightforward consequence of the implicit function theorem (see Wiggins 1988).

Hence within $O(\epsilon)$ of the zero sets of the Melnikov function which satisfy the above derivative criterion are intersection manifolds of $W^s(\tau_\epsilon^a)$ and $W^u(\tau_\epsilon^b)$. We refer to this class of manifold intersections as *primary intersection manifolds* (PIMS), and these manifolds are central to the upcoming definition of a *lobe*. We approximately portray (i.e. to $O(\epsilon)$) some PIMS by plotting the zero sets of a two frequency Melnikov function of the form

$$M(s, \theta_1, \theta_{2_0} = 0; \nu) = A_0(\mu) + A_1(\mu, \omega_1) \sin(\omega_1 s + \theta_1) + A_2(\mu, \omega_2) \sin(\omega_2 s). \quad (2.10)$$

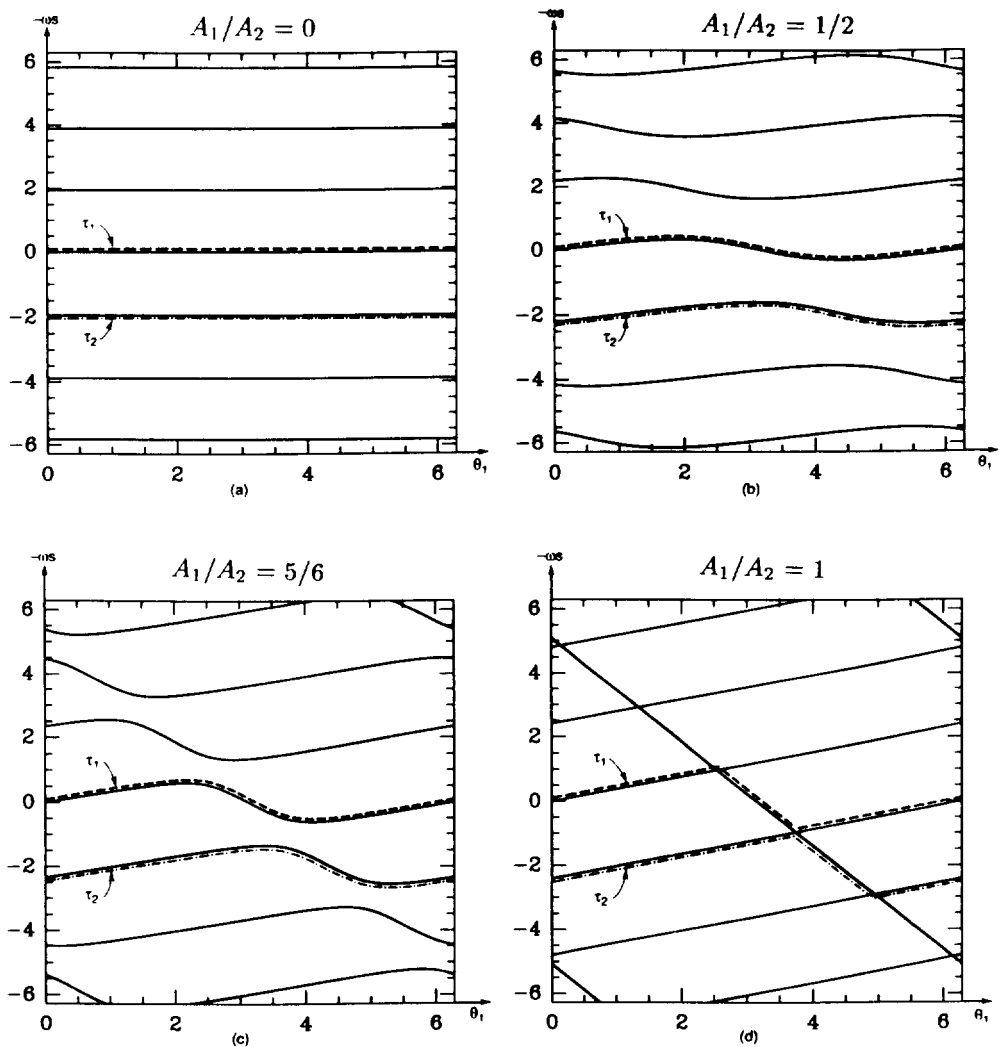


Figure 4. Some zero sets of the Melnikov function given by equation (2.10), for a range of parameter values ($\omega_1 = g\omega_2 \equiv \omega$, except for (h) where $\omega_1 = \omega_2/2 \equiv \omega$). The broken and chain lines represent two choices of individual PIMs, τ_1 and τ_2 .

Figures 4(a) to (g) show the zero sets near $s = 0$ for a range of A_1/A_2 values and $A_0 = 0$, $\omega_2/\omega_1 = g^{-1}$, where g is the golden mean $((\sqrt{5} - 1)/2)$. For $A_1/A_2 < 1$ the PIMs are non-intersecting 1-tori, for $A_1/A_2 > 1$ they are non-intersecting spirals, and for $A_1/A_2 = 1$ they are intersecting spirals (or equivalently intersecting 1-tori). For other ratios of ω_2/ω_1 (both commensurate and incommensurate), the geometry of the PIMs as A_1/A_2 is varied (keeping $A_0 = 0$) is qualitatively similar, as will become more apparent with some additional figures in section 3. There is a slightly technical point about the non-generic case of $A_1/A_2 = 1$: that the zero sets cross does not necessarily imply that the PIMs cross; however, we have shown that for each crossing of the $M = 0$ sets the intersection manifolds cross within $O(\varepsilon)$, and for the ovp flow each such crossing occurs for the same value $A_1/A_2 = 1 + O(\varepsilon)$ (see Beigie et al

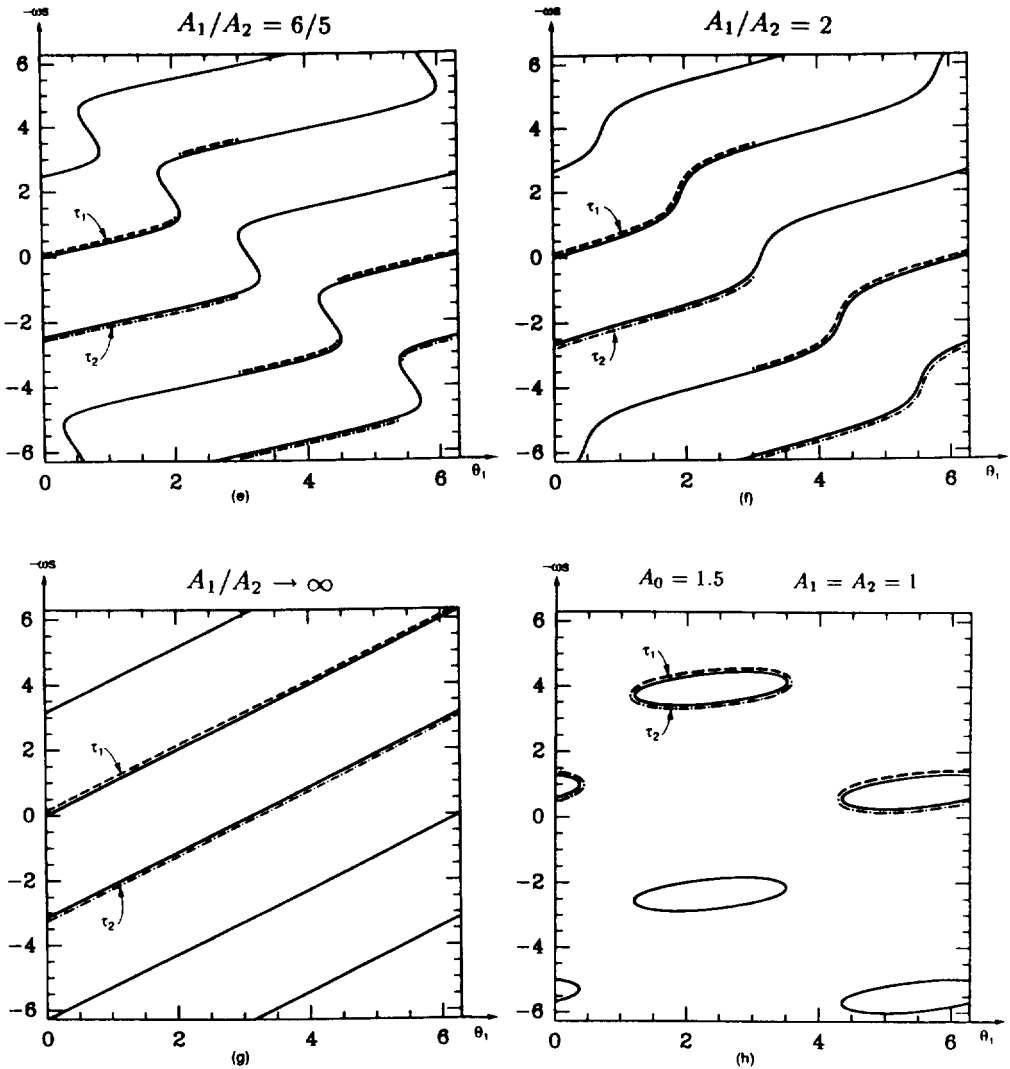


Figure 4—(continued)

1991b). Figure 4(h) shows a case with $A_0 = \pm 1.5$, $A_1 = A_2 = \pm 1$, and $\omega_2/\omega_1 = 2$: the intersection manifolds exist in a *subset* of $[0, 2\pi)$. As will be seen in an upcoming theorem, the geometry of intersection manifolds for $A_0 \neq 0$ may be qualitatively different for commensurate and incommensurate frequencies, since it is in the commensurate frequency case alone that for certain parameter values the intersection manifolds exist on a subset of $[0, 2\pi)$. Figure 4 also shows two choices of *individual* PIMS τ_1 and τ_2 . In Beigie *et al* (1991b) we provide an explicit definition of individual PIMS and study their geometries; their nature should be clear from the figures and we will not go into details here. The essential point is that they are $(l-1)$ -dimensional piecewise continuous intersection manifolds which intersect each *phase slice* $\chi(\bar{\theta}) \equiv \{(x, \theta) \mid \theta = \bar{\theta} \in T^{l-1}\}$ in a single point, for each $\bar{\theta}$ in some subset of T^{l-1} , denoted \mathcal{X}^{l-1} . Typically this subset is *identical* to T^{l-1} , but it does not have

to be, as in figure 4(h). The motivation for allowing the PIMS to be piecewise continuous will become clear momentarily. The central result from our study of the geometries of PIMS via the Melnikov function is given below.

Theorem 2.3. Suppose there exists a point $(\bar{s}, \bar{\theta}_1, \dots, \bar{\theta}_{l-1})$ for some \bar{v} such that

- (1) $M(\bar{s}, \bar{\theta}_1, \dots, \bar{\theta}_{l-1}, \bar{\theta}_l; \bar{v}) = 0$,
- (2) $(\partial M / \partial s)(\bar{s}, \bar{\theta}_1, \dots, \bar{\theta}_{l-1}, \bar{\theta}_l; \bar{v}) \neq 0$.

If the frequencies are all mutually incommensurate, then $W^s(\tau_\epsilon^a)$ intersects $W^u(\tau_\epsilon^b)$ in a countable infinity of $(l-1)$ -dimensional surfaces that can be represented as graphs over T^{l-1} . If one or more pairs of frequencies are commensurate, then $W^s(\tau_\epsilon^a)$ intersects $W^u(\tau_\epsilon^b)$ in a countable infinity of $(l-1)$ -dimensional surfaces that can be represented as graphs over all T^{l-1} or some subset of T^{l-1} , denoted \mathcal{X}^{l-1} .

Proof. See Beigie et al (1991b).

With the PIMS we can now define and study an invariant lobe structure. To deal with the arbitrary geometries of $W^s(\tau_\epsilon^a)$ and $W^u(\tau_\epsilon^b)$, it will prove convenient to define a *lobe* as an appropriate union over phase slices of a simply connected two dimensional region in each phase slice. The first two of the following definitions thus pertain to phase slices.

Definition 2.1. For $\bar{\theta} \in \mathcal{X}^{l-1}$, the intersection of the PIM τ with the phase slice $\chi(\bar{\theta}) \equiv \{(x, \theta) \mid \theta = \bar{\theta}\}$ defines a unique point. We refer to this point $p(\bar{\theta}) = \tau \cap \chi(\bar{\theta})$ as a *primary intersection point* (PIP).

Definition 2.2. Let $p_1(\bar{\theta})$ and $p_2(\bar{\theta})$ denote two PIPs in the phase slice $\chi(\bar{\theta})$. Then let $S[p_1(\bar{\theta}), p_2(\bar{\theta})]$ and $U[p_1(\bar{\theta}), p_2(\bar{\theta})]$ denote the segments of $W^s(\tau_\epsilon^a) \cap \chi(\bar{\theta})$ and $W^u(\tau_\epsilon^b) \cap \chi(\bar{\theta})$, respectively, from $p_1(\bar{\theta})$ to $p_2(\bar{\theta})$. We say the two PIPs are *adjacent* if $S[p_1(\bar{\theta}), p_2(\bar{\theta})]$ and $U[p_1(\bar{\theta}), p_2(\bar{\theta})]$ contain no other PIPs than $p_1(\bar{\theta})$ and $p_2(\bar{\theta})$.

Definition 2.3. For all $\bar{\theta} \in \mathcal{X}^{l-1}$, let $p_1(\bar{\theta}), p_2(\bar{\theta})$ denote adjacent PIPs in the phase slice $\chi(\bar{\theta})$. A *lobe* \mathcal{L} is an $(l+1)$ -dimensional region in Σ^{θ_0} such that:

- (a) for all $\bar{\theta} \in \mathcal{X}^{l-1}$, $\mathcal{L} \cap \chi(\bar{\theta})$ is the region in $\chi(\bar{\theta})$ bounded by $S[p_1(\bar{\theta}), p_2(\bar{\theta})] \cup U[p_1(\bar{\theta}), p_2(\bar{\theta})]$, and
- (b) the sign of $M(s, \bar{\theta}_1, \dots, \bar{\theta}_{l-1}, \theta_l; v)$ in the interval $s(\bar{\theta}) \in [s_1(\bar{\theta}), s_2(\bar{\theta})]$, where $s_i(\bar{\theta})$ is the s value of $p_i(\bar{\theta})$, for $i = 1, 2$, is independent of $\bar{\theta} \in \mathcal{X}^{l-1}$.

The motivation for condition (b) should be clear from the familiar two-dimensional lobes associated with time-periodic vector fields: there are two classes of lobes, as distinguished by the sign of the Melnikov function associated with the lobe. When we form the union of regions over different phase slices, we want to choose regions from the same class.

This general definition of a lobe, as with the definition of a PIM, allows several geometrical possibilities; however, our understanding of the geometry of PIMS translates into an understanding of the geometry of lobes. Indeed, for the examples in figure 4, let us choose, for all $\bar{\theta}_1 \in \mathcal{X}^1$, $p_1(\bar{\theta}_1) = \tau_1 \cap \chi(\bar{\theta}_1)$ and $p_2(\bar{\theta}_1) = \tau_2 \cap \chi(\bar{\theta}_1)$. From the PIMS in figure 4 one can visualize a lobe by remembering that its 'thickness', i.e. the separation of $W^s(\tau_\epsilon^a)$ and $W^u(\tau_\epsilon^b)$ normal to $W^s(\tau_\epsilon^a) \cap W^u(\tau_\epsilon^b)$, comes out of the page (see figure 5). For the toral PIMS of figures 4(a), (b), (c), the

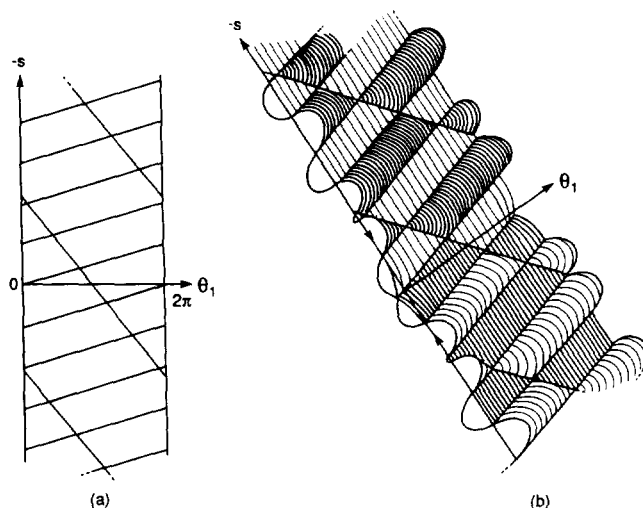


Figure 5. Visualizing lobes by showing the suppressed dimension.

three-dimensional lobe is a simply connected region whose boundary divides Σ^{θ_0} into an inside and outside; for the spiral PIMS of figure 4(g), the lobe has a discontinuity at $\theta_1 = 0$; for the intersecting PIMS of figure 4(d) the lobe has no discontinuity but 'pinches off' to zero volume at an isolated θ_1 value; for the PIMS of figure 4(e) and (f) with one or more discontinuities in $[0, 2\pi)$, the lobe has discontinuities in $[0, 2\pi)$; for the PIMS of figure 4(h) which contain gaps in T^1 , the lobes contain gaps in T^1 .

From the definition of lobes and theorem 2.3, when the system satisfies the conditions of theorem 2.3 then $W^s(\tau_e^a)$ and $W^u(\tau_e^b)$ form the boundaries of a countable infinity of $(l+1)$ -dimensional lobes in Σ^{θ_0} , and hence we refer to $W^s(\tau_e^a)$ and $W^u(\tau_e^b)$ as the boundary of an *invariant lobe structure* in Σ^{θ_0} (see the top part of figure 6; the remaining parts of this figure will become clear as we proceed).

2.2. Deriving a sequence of two-dimensional time-dependent lobe structures from the invariant $(l+1)$ -dimensional lobe structure

2.2.1. The autonomous system. Our eventual goal is to study the dynamics of the non-autonomous system (2.2) under the sequence of maps $\{T_e(\cdot; n); n \in \mathbb{Z}\}$. From expression (2.4), and the comments that followed, the study of trajectories of (2.2) through x_0 at time $t=0$ for all x_0 is equivalent to the study of the x component of the trajectories of the autonomous system (2.3) through $(x_0, \theta_{10}, \dots, \theta_{l0})$ at time $t=0$ for all x_0 . The coordinates of the solution $\Phi(t, t_0=0, x_0, \theta_{10}, \dots, \theta_{l0})$ at time $t = (2\pi/\omega_l)n$ ($n \in \mathbb{Z}$) for all x_0 lie in the plane

$$\left\{ (x, \theta_1, \dots, \theta_l) \mid \theta_i = \theta_{i0} + 2\pi \frac{\omega_i}{\omega_l} n; i = 1, \dots, l \right\}.$$

We shall refer to this plane as the n th *time slice* of the phase space of the autonomous system, which is of course identical to a phase slice. To study the dynamics of (2.2), we are interested in the sequence of time slices of the autonomous system phase space, and the intersection of the invariant lobe structure

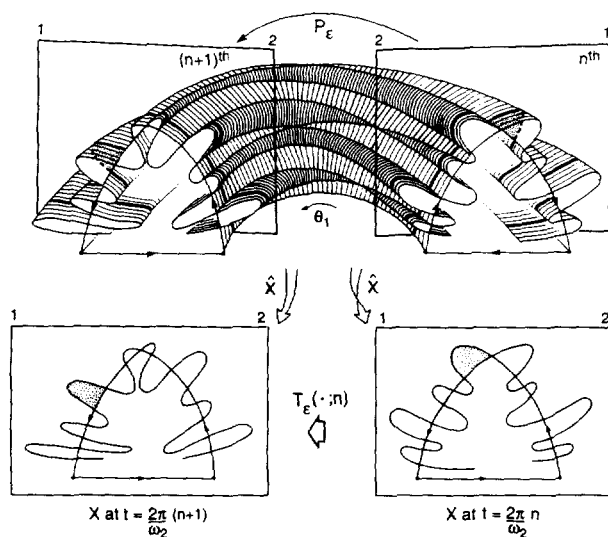


Figure 6. The boundaries of a three-dimensional invariant lobe structure in $\Sigma^{\theta_{10}}$ for a two frequency heteroclinic case (this geometry corresponds to the top half of the OVP flow). The n th and $(n+1)$ th time slice of $\Sigma^{\theta_{10}}$ are shown, along with the corresponding projections onto X to obtain the time-dependent two-dimensional lobe structures in the non-autonomous system phase space. A lobe mapping between successive time slices is portrayed.

with these time slices defines a sequence of two-dimensional time-dependent lobe structures. For the single frequency case, each time slice is identical to the plane defined by the Poincaré section $\Sigma^{\theta_{10}}$. Hence a two-dimensional lobe structure in this plane is immediately provided by the invariant lobe boundaries $W^s(\tau_\epsilon^a) \cup W^u(\tau_\epsilon^b)$ in $\Sigma^{\theta_{10}}$ (here the τ s refer to points). For the two frequency case, the n th time slice is a plane in the three-dimensional Poincaré section $\Sigma^{\theta_{10}}$ defined by

$$\chi\left(\theta_{10} + 2\pi \frac{\omega_1}{\omega_2} n\right) = \left\{ (x, \theta_1) \mid \theta_1 = \theta_{10} + 2\pi \frac{\omega_1}{\omega_2} n \right\}$$

(see figure 6). For $\theta_{10} + 2\pi(\omega_1/\omega_2)n \in \mathcal{I}^1$ the intersection of this plane with the invariant lobe boundaries $W^s(\tau_\epsilon^a) \cup W^u(\tau_\epsilon^b)$ in $\Sigma^{\theta_{10}}$ defines the boundary of a two-dimensional lobe structure:

$$\Gamma\left(\theta_{10} + 2\pi \frac{\omega_1}{\omega_2} n\right) = \{W^s(\tau_\epsilon^a) \cup W^u(\tau_\epsilon^b)\} \cap \chi\left(\theta_{10} + 2\pi \frac{\omega_1}{\omega_2} n\right)$$

(we henceforth assume that $\theta_{10} \in \mathcal{I}^1$, i.e. that the initial time slice contains a lobe structure, and hence, by invariance under P_ϵ , all time slices have a lobe structure). The intersection of the set of PIMS with the plane $\chi(\theta_{10} + 2\pi(\omega_1/\omega_2)n)$ defines a countable infinity of PIMS, and a countable infinity of two-dimensional lobes are defined by regions in $\chi(\theta_{10} + 2\pi(\omega_1/\omega_2)n)$ bounded by segments of stable and unstable manifolds in that plane that connect adjacent PIMS. The geometry of the two-dimensional lobe structure in each time slice is thus the same as the familiar invariant lobe structure of the time-periodic case. The two frequency case easily extends to the l frequency case by replacing θ_1 , ω_1 , and θ_{10} with $\theta \equiv (\theta_1, \theta_2, \dots, \theta_{(l-1)})$, $\omega \equiv (\omega_1, \omega_2, \dots, \omega_{(l-1)})$, and $\theta_0 \equiv (\theta_{10}, \theta_{20}, \dots, \theta_{(l-1)0})$,

respectively. We will frequently refer to the expression $\theta_0 + 2\pi(\omega/\omega_l)n$, so it will be useful to remember that it specifies along with θ_{l_0} the phases of the perturbation at the n th time slice.

2.2.2. The non-autonomous system. The sequence of two-dimensional lobe structures defined in the autonomous system phase space will suffice to study the trajectories of (2.2). Indeed some might prefer to remain in the autonomous system phase space since, after all, the invariant manifolds that constrain the motion are in this space. However, when we study the motion of a two-dimensional non-autonomous physical system of the form (2.2), there is conceptual appeal in returning to the physical space, X , rather than dealing with time slices of an $(l+2)$ -dimensional phase space. This is a simple matter of projecting the coordinates of the autonomous solution $\Phi(t, t_0=0, x_0, \theta_{1_0}, \theta_{2_0}, \dots, \theta_{l_0})$ onto X , since the x component of this solution solves the non-autonomous system (2.2). Let \hat{X} denote the projection operator onto X :

$$\begin{aligned}\hat{X} : X \times T^l &\mapsto X \\ (x, \theta_1, \theta_2, \dots, \theta_l) &\mapsto x.\end{aligned}$$

Then the boundaries of the two-dimensional lobe structure in X at time $t = (2\pi/\omega_l)n$ are given by

$$\gamma(n) = \hat{X}\left(\Gamma\left(\theta_0 + 2\pi\frac{\omega}{\omega_l}n\right)\right)$$

as shown in figure 6. We stress at the outset that for the sake of simple notation, when describing the projected quantities here and elsewhere, the dependence on the initial phases (θ_0, θ_{l_0}) and frequencies (ω, ω_l) is to be understood implicitly. Some PIPS and two-dimensional lobes are shown explicitly in figure 7. Earlier (following theorem 2.2) we referred to one class of intersection manifolds (those continued from the $M=0$ sets—the PIMS), and in figure 7 we explicitly see the two types of intersection sets: PIPS, marked by crosses, which are continued from, and hence within $O(\varepsilon)$ of, the zero sets of the Melnikov function, and SIPS (secondary intersection points), marked by solid dots, which need not be within $O(\varepsilon)$ of the zero sets. Similarly, in $\Sigma^{\theta_{l_0}}$ one has, in addition to PIMS, secondary intersection manifolds (SIMS). We refer to the time-dependent lobe structure in X as a *homoclinic or heteroclinic tangle region*. We proceed to study the dynamics in this region by using the invariance under P_ε of the higher-dimensional lobe structure in $\Sigma^{\theta_{l_0}}$. Note that, since in the single frequency case the time slices and hence the lobe boundaries in each slice are the same for all values of n , the projected lobe structure on X recovers the familiar invariant lobe structure of the time-periodic case.

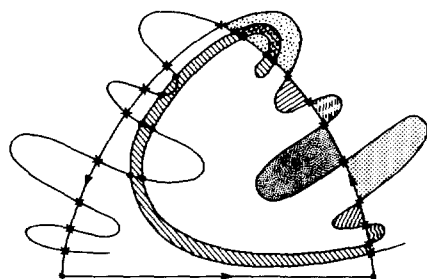


Figure 7. Showing explicitly some intersection points and two-dimensional lobes in X for a given time sample. The PIPS are marked by crosses, and the SIPS are marked by dots. Each lobe is filled in differently.

3. Describing transport in the tangle regions

We proceed to describe the global dynamics within the homoclinic and heteroclinic tangle regions of (2.2). It is a straightforward consequence of the invariance of $W^s(\tau_\epsilon^a)$ and $W^u(\tau_\epsilon^b)$ (see Beigie et al 1991b) that each lobe in the n th time slice maps under P_ϵ to a lobe in the $(n+1)$ th time slice; equivalently, projecting onto X , each lobe in X at time $t = (2\pi/\omega_l)n$ maps under $T_\epsilon(\cdot; n)$ to a lobe in X at time $t = (2\pi/\omega_l)(n+1)$ (refer back to figure 6). The mapping of one lobe to another under successive time samples provides a global picture of the dynamics, i.e. of *transport in phase space*, in the tangle region (see figure 8). We concentrate in this section on describing transport in terms of two-dimensional lobes mapping from one to another. We facilitate this discussion with the following definition.

Definition 3.1. We refer to the adjacent PIPs that are used to define a two-dimensional lobe in a phase slice as the *bounding* PIPs of that lobe. Further, we say a lobe is *between* two arbitrary PIPs in a phase slice if its bounding PIPs lie on or between these two PIPs (relative to the curve defined by either the stable or unstable manifold in that phase slice).

3.1. Transport in the time-periodic case

Before discussing transport in the $l \geq 2$ case, let us summarize the results of the single frequency case, illustrated by the OVP flow (see the appendix). Figure 9 shows the invariant two-dimensional lobe structure in the $x_2 \geq 0$ region of $\Sigma^{\theta_{l_0}=0}$. An upper core boundary, shown by the broken dotted line, is defined by

$$C = U[q_\epsilon^b, p_c] \cup S[p_c, q_\epsilon^a] \cup x_1[q_\epsilon^a, q_\epsilon^b]$$

where p_c denotes the PIP at $s=0$ and $x_1[q_\epsilon^a, q_\epsilon^b]$ denotes the segment along the x_1 axis from the left fixed point q_ϵ^a to the right fixed point q_ϵ^b (the meaning of $U[q_\epsilon^b, p_c]$ and $S[p_c, q_\epsilon^a]$ should be clear from section 2). One can formally choose any PIP to play the role of p_c in defining the core; however, p_c is chosen here to give a core that best approximates the unperturbed core. Regions in $\Sigma^{\theta_{l_0}=0}$ are *entrained into/detrained from* the core if they are mapped from outside/inside the core to inside/outside the core. The lobes between $P_\epsilon^{-m}(p_c)$ and $P_\epsilon^{-(m+1)}(p_c)$ are mapped under P_ϵ to the lobes between $P_\epsilon^{-(m-1)}(p_c)$ and $P_\epsilon^{-m}(p_c)$ such that they preserve orientation and relative ordering (this preservation is a consequence of the fact that the Poincaré map is obtained by discrete time sampling of trajectories of ODES, as discussed in Wiggins (1988)). By the definition of the core boundary, the lobes between p_c and $P_\epsilon^{-1}(p_c)$ reverse their orientation relative to the core boundary under application of P_ϵ ; all

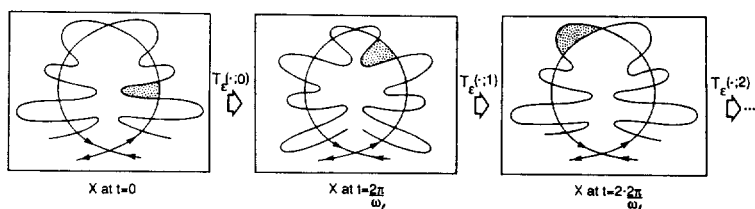


Figure 8. A lobe mapping within the sequence of lobe structures in X for an l frequency homoclinic case.

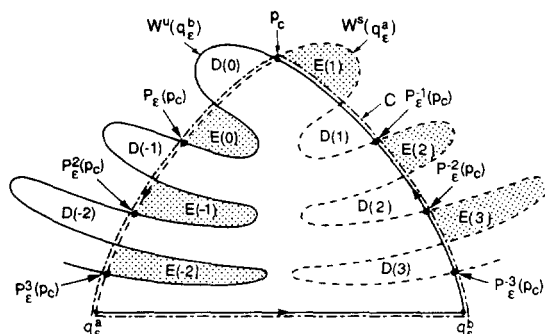


Figure 9. The invariant lobe structure in $\Sigma^{010=0}$ for the top half of the time-periodic ovr flow.

other lobes preserve under P_ϵ their orientation relative to the core boundary. Borrowing the terminology from MacKay *et al* (1984), we refer to the lobes between p_c and $P_\epsilon^{-1}(p_c)$ as *turnstile lobes*. We assume that each turnstile lobe and its image is entirely inside or entirely outside the core (which is usually the case), with the understanding that if any one of the lobes has pieces both inside and outside the core, one can redefine the lobes to be the pieces of the old lobes contained in the appropriate region (inside or outside the core). Each turnstile lobe is either entrained into or detrained from the core under P_ϵ ; these lobes are the *only* mechanism for transport inside and outside the core under P_ϵ . This motivates the following definition of *sets* of lobes: for $m \in \mathbb{Z}$, $m > 0$, $E(m)$ and $D(m)$, respectively, denote the sets of lobes that are entrained into/detrained from the core upon the m th iterate of P_ϵ ; for $m \leq 0$ they denote the sets of lobes that are detrained from/entrained into the core upon the $(|m| + 1)$ th iterate of P_ϵ^{-1} . We refer to $E(m)$ and $D(m)$ as sets of *entraining* and *detraining* lobes, respectively. From previous remarks, $E(m) \cup D(m)$ is the set of lobes between $P_\epsilon^{-(m-1)}(p_c)$ and $P_\epsilon^{-m}(p_c)$. For the single frequency ovr flow studied by Rom-Kedar *et al* (1990), $E(m)$ and $D(m)$ each contain one lobe for all $m \in \mathbb{Z}$. For single frequency forcing the number of lobes is independent of m and (generically) of whether the lobes are entraining or detraining. In near-integrable systems, the Melnikov function can be used to determine the number of lobes in each set $E(m)$ and $D(m)$; otherwise one must numerically compute the lobe structure. We will discuss these procedures in the multiple frequency case, which will also cover the single frequency case. By their definition the sets of lobes map under P_ϵ according to

$$P_\epsilon(L(m)) = L(m - 1) \quad \forall m \in \mathbb{Z} \quad (3.1)$$

where $L = E$ or D . Since the sets of lobes in this case are identical to individual lobes, equation (3.1) specifies the dynamics of individual lobes. In our discussion of the multiple frequency case we will explain the more complicated scenario when the sets contain more than one lobe. Since the Poincaré section is two-dimensional, it is a trivial matter to project onto the 'physical' space X , so we will not bother to do it here.

3.2. Transport in the quasiperiodic case

We describe transport in a manner that is step by step analogous to the approach in the time-periodic case. We illustrate our discussion with two frequency near-

integrable systems with a Melnikov function of the form

$$M(s, \theta_1, \theta_{2_0} = 0; \nu) = A_1(\mu, \omega_1) \sin(\omega_1 s + \theta_1) + A_2(\mu, \omega_2) \sin(\omega_2 s) \quad (3.2)$$

(refer to section 2 and the appendix). We stress, though, that the description of transport does not rely on near-integrability. To illustrate transport using a three-dimensional portrayal of the lobe structure in $\Sigma^{\theta_0=0}$ would certainly be cumbersome; we thus suppress the dimension normal to the unperturbed homoclinic/heteroclinic manifold, as done when we plotted the zero sets of the Melnikov function in the previous section. As described in that section, the set of $M = 0$ curves portray to $O(\varepsilon)$ the set of PIMS, and hence portray the dividing lines between invariant three-dimensional lobes, whose thickness comes out of the page (refer back to figure 5). Consider some simple choices of A_i and ω_i shown in figures 10(a)–(g). The shading and labelling of these plots will be explained as our notation is developed. We include commensurate frequencies along with incommensurate ones: even though the commensurate cases are really periodic, they are still illuminating for our two frequency prescription. Indeed, for many commensurate frequency cases, the common period of the two frequencies is much longer than the individual periods, so there is motivation to sample at one of the two periods and hence use the two frequency formalism.

We first define an invariant core in Σ^{θ_0} , which necessitates the definition of a manifold that plays the role that p_c does in the time-periodic case. Let τ_c be an $(l-1)$ -dimensional PIM.

Definition 3.2. A homoclinic core boundary \mathcal{C} in Σ^{θ_0} is the l -dimensional surface

$$\{U[p_\varepsilon(\bar{\theta}), p_c(\bar{\theta})] \cup S[p_c(\bar{\theta}), p_\varepsilon(\bar{\theta})] \mid \bar{\theta} \in \mathcal{X}^{l-1}\},$$

where $p_\varepsilon(\bar{\theta}) = \tau_\varepsilon \cap \chi(\bar{\theta})$ and $p_c(\bar{\theta}) = \tau_c \cap \chi(\bar{\theta})$.

The heteroclinic core boundary involves two heteroclinic orbits, and hence has more than one possible form. We will worry about only one of the heteroclinic orbits and hence make the following definition.

Definition 3.3. A heteroclinic core boundary \mathcal{C} in Σ^{θ_0} is the l -dimensional surface

$$\{U[p_\varepsilon^b(\bar{\theta}), p_c(\bar{\theta})] \cup S[p_c(\bar{\theta}), p_\varepsilon^a(\bar{\theta})] \cup O[p_\varepsilon^a(\bar{\theta}), p_\varepsilon^b(\bar{\theta})] \mid \bar{\theta} \in \mathcal{X}^{l-1}\},$$

where $O[p_\varepsilon^a(\bar{\theta}), p_\varepsilon^b(\bar{\theta})]$ denotes the segment of the core boundary in the phase slice $\chi(\bar{\theta})$ due to the other heteroclinic orbit, which may be the same as the unperturbed orbit (as in the ovp flow), or may involve another heteroclinic tangle.

As done in the time-periodic case with p_c , we specialize our concern to a specific choice of τ_c , and hence of the core:

$$\tau_c = \{(x, \theta) \mid x = \zeta(\theta), \theta \in \mathcal{X}^{l-1}\}$$

where $\zeta(\theta = \bar{\theta})$ is the x value of the PIP in the phase slice $\chi(\bar{\theta})$ such that $\mathcal{C} \cap \chi(\bar{\theta})$ best approximates the unperturbed core in the phase slice $\chi(\bar{\theta})$. For example, the unperturbed heteroclinic orbits in X of the ovp flow are symmetric about their spatial midpoint ($s \equiv 0$), and a natural choice for τ_c is for $\zeta(\theta = \bar{\theta})$ to be the x value of the PIP in the $\chi(\bar{\theta})$ plane whose s value is closest to zero. Figure 10 shows this choice of τ_c for the two frequency examples: for 10(c), (d) and (g), τ_c is a 1-torus,

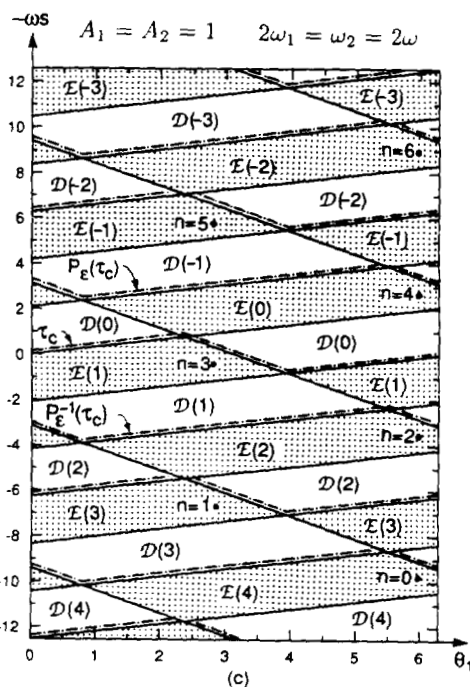
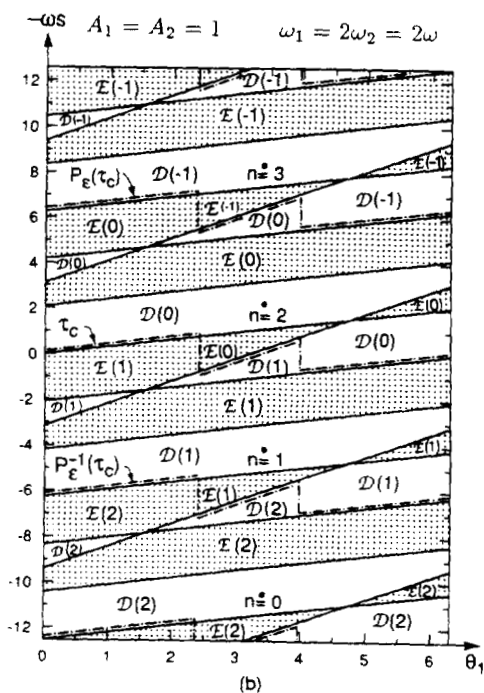
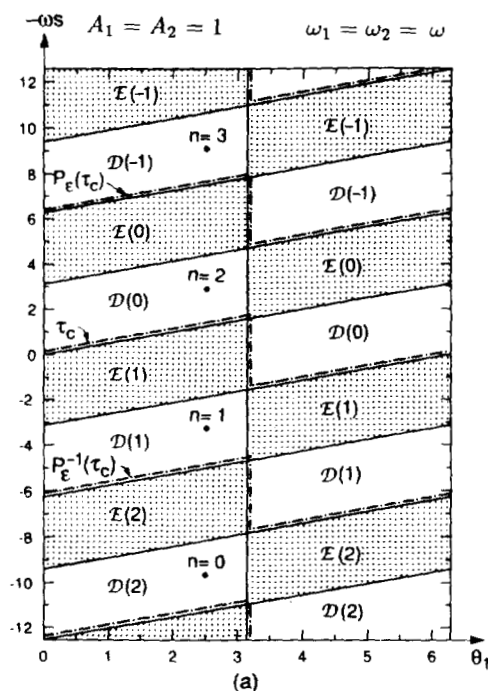


Figure 10. Portrayal and labelling, in (s, θ_1) space, of the invariant three-dimensional lobes, PIMS, τ_c and $P_\epsilon^{\pm 1}(\tau_c)$ for systems with Melnikov function of the form (3.2).

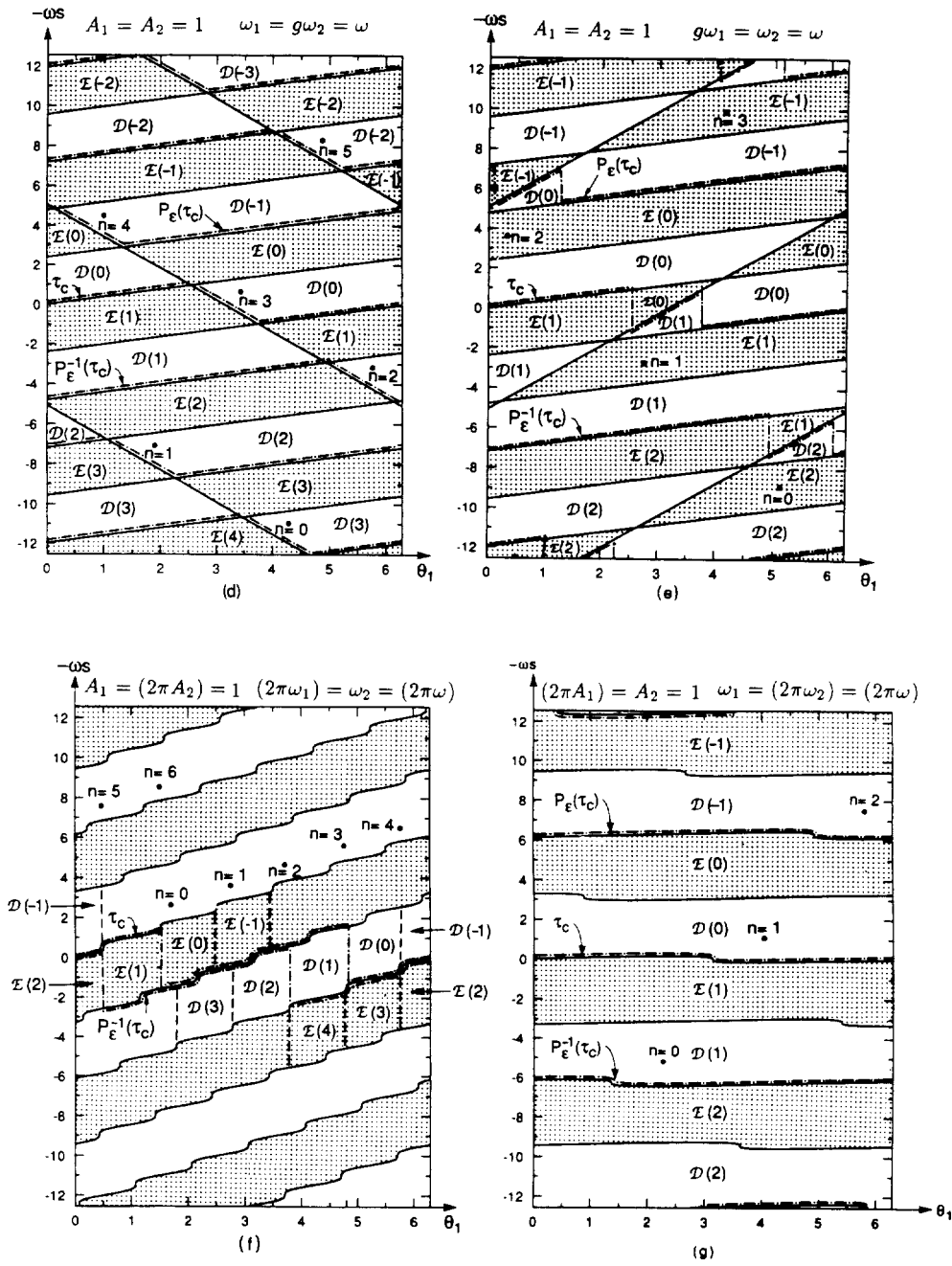


Figure 10—(continued)

while in 10(a), (b), (e) and (f) it is a piecewise continuous set of segments of spirals. Note that, for visual clarity, the piecewise continuous sets are joined by vertical broken lines at the points of discontinuity; it should be understood that τ_c is really single-valued everywhere. The invariant core for this choice of τ_c immediately defines a two-dimensional core in any phase slice, $\chi(\bar{\theta})$, $\bar{\theta} \in \mathcal{X}^{t-1}$, and equivalently in X for any time sample:

$$C(\bar{\theta}) = \mathcal{C} \cap \chi(\bar{\theta})$$

$$c(n) = \hat{X} \left(C \left(\theta_0 + 2\pi \frac{\omega}{\omega_l} n \right) \right). \quad (3.3)$$

We stress two fundamental points regarding the invariant core and the resulting *core sequence* $\{c(n); n \in \mathbb{Z}\}$. First, since the invariant core boundary in general may be piecewise continuous in θ , it does not necessarily divide the Poincaré section Σ^{θ_0} into an inside and outside. This would present a problem for more general systems, such as the k degree-of-freedom systems discussed in Wiggins (1990). *However, due to the trivial nature of the dynamics in θ (everything maps from phase slice to phase slice), all that is needed of the invariant core boundary is that it divides each phase slice into an inside and an outside, which our piecewise continuous core boundary indeed does.* By definition, for each n , $C(\theta_0 + 2\pi(\omega/\omega_l)n)$ divides its two-dimensional time slice into an inside and an outside, and $c(n)$ divides X into an inside and outside. Second, though τ_c and hence the invariant core boundary may in certain cases seem complicated and geometrically unappealing, by its very definition it ensures that for each n the core $c(n)$ most closely resembles the unperturbed core boundary. The aim then in the definition of τ_c is geometric appeal in each phase slice.

We now define entrainment and detrainment with respect to the core sequence.

Definition 3.4. The regions in the n th time slice $\chi(\theta_0 + 2\pi(\omega/\omega_l)n)$ of Σ^{θ_0} that are *entrained into/detrained from the core under P_e* are the regions in that slice that are outside/inside $C(\theta_0 + 2\pi(\omega/\omega_l)n)$ and are mapped to inside/outside $C(\theta_0 + 2\pi(\omega/\omega_l)(n+1))$ in the $(n+1)$ th time slice. Equivalently, the regions in X that are *entrained into/detrained from the core under $T_e(\cdot; n)$* are the regions in X that are outside/inside $c(n)$ and are mapped to inside/outside $c(n+1)$.

We have defined entrainment and detrainment with respect to a core that changes shape and area from one time sample to the next, which is markedly different from the time-periodic case, where the core corresponds to a fixed spatial region. Thus, before continuing we should motivate such a definition. Imagine for example a two-dimensional material blob that pulsates in some quasiperiodic fashion, as shown in the sequence of snapshots in figure 11(a). We do not want to say there is any entrainment or detrainment here. Though we could define a fixed spatial core and monitor the back and forth motion across this core due to the pulsation, we prefer to define a time-dependent core boundary as the material blob boundary, and say that there is no entrainment/detrainment with respect to this pulsating core. Now consider a material blob which, on top of a net pulsation, shows lobe formations that one would like to associate with entrainment/detrainment, as shown in figure 11(b). Again, here our preferred way to monitor entrainment/detrainment is with respect to a pulsating core boundary, as shown by the broken line in figure 11(b), since this prevents us from counting net pulsation as

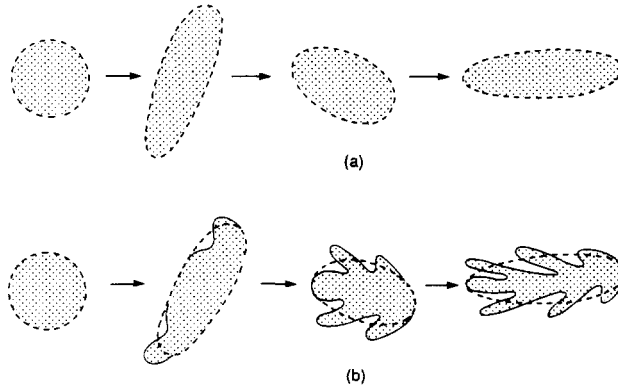


Figure 11. (a) A pulsating material blob and (b) a pulsating and lobe-forming material blob. A defined core boundary is shown by the broken line.

entrainment/detrainment. The obvious complication is: how do we distinguish between ‘overall pulsation’ and ‘lobe formation’? When a material blob deforms from one time sample to the next, to distinguish between the ‘reversible’ pulsation deformations and the ‘irreversible’ lobe-like deformations, we need to know the long-time behaviour of the material (see figure 12(a)). An essential feature of our lobe structure comes into play here: the invariance under P_ϵ of $W^s(\tau_\epsilon^a)$ and $W^u(\tau_\epsilon^b)$ implies that these manifolds contain the long term information needed to provide a natural way of distinguishing between net pulsation and lobe formation for system (2.2). For example, returning to the ovp flow, in the region in X at time $t = (2\pi/\omega_2)(n+1)$ where $T_\epsilon(c(n); n)$ shows lobe-like deformations, the curve $\hat{X}(W^s(\tau_\epsilon^a) \cap \chi(\theta_{l_0} + 2\pi(\omega_1/\omega_2)(n+1)))$ provides the natural dividing line between ‘inside’ and ‘outside’ since by definition the points on $W^s(\tau_\epsilon^a)$ asymptote to τ_ϵ^a (see

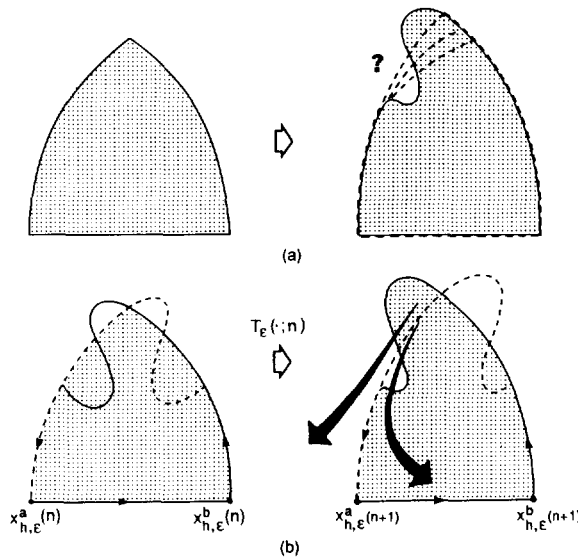


Figure 12. (a) A material blob deforms; the broken lines show possible dividing lines between ‘inside’ and ‘outside’. (b) The stable manifold (the broken line) provides a natural dividing line where the lobes form between ‘inside’ and ‘outside’.

figure 12(b)); fluid particles immediately to the left of the highlighted $\hat{X}(W^s(\tau_\epsilon^a) \cap \chi(\theta_{1_0} + 2\pi(\omega_1/\omega_2)(n+1)))$ segment are destined to move further and further in the $-x_1$ direction (their x_1 component eventually going to $-\infty$), while fluid particles immediately to the right are destined to move further and further in the $+x_1$ direction until they are rotated around the core to be either detrained or rotated around the core again, and so on. The invariant manifolds $W^s(\tau_\epsilon^a)$, $W^u(\tau_\epsilon^b)$ thus provide a natural way to define a time-dependent core in X with which to monitor entrainment/detrainment for the sequence of non-autonomous maps $\{T_\epsilon(\cdot; n); n \in \mathbb{Z}\}$. Though for concreteness we have phrased our motivation in the context of a fluid, we do not wish the reader to lose sight of the significance of a time-dependent core in relation to the fundamental notion of bounded and unbounded motion in non-integrable systems. Indeed points within the core may escape, and hence their motion may not be bounded; however, we shall soon see that they can escape only via the detraining turnstile lobes, and hence, when detrainment leads to unbounded motion (as is the case in the OVP flow), it is meaningful to regard the time-dependent core boundary as the *last frontier* between bounded and unbounded motion and to view the points within the core as bounded until they escape via the detraining turnstile lobes to unbounded motion.

With a motivated definition of a core sequence, we address the global dynamics associated with lobes mapping from one time slice to the next. The two-dimensional lobes in $\chi(\bar{\theta})$, $\bar{\theta} \in \mathcal{X}^{l-1}$, between $P_\epsilon^{-m}(\tau_c) \cap \chi(\bar{\theta})$ and $P_\epsilon^{-(m+1)}(\tau_c) \cap \chi(\bar{\theta})$ are mapped under P_ϵ to the lobes in $\chi(\bar{\theta} + 2\pi(\omega/\omega_l))$ between $P_\epsilon^{-(m-1)}(\tau_c) \cap \chi(\bar{\theta} + 2\pi(\omega/\omega_l))$ and $P_\epsilon^{-m}(\tau_c) \cap \chi(\bar{\theta} + 2\pi(\omega/\omega_l))$ such that they preserve their orientation and relative ordering (this preservation is again a consequence of the fact that P_ϵ is obtained by discrete time sampling of the trajectories of ODES). By the definition of the invariant core boundary, the lobes in $\chi(\bar{\theta})$ between $\tau_c \cap \chi(\bar{\theta})$ and $P_\epsilon^{-1}(\tau_c) \cap \chi(\bar{\theta})$ reverse their orientation relative to the core boundary under the application of P_ϵ ; all other lobes in that phase slice preserve under P_ϵ their orientation relative to the core boundary. Hence we refer to the lobes in $\chi(\bar{\theta})$ between $\tau_c \cap \chi(\bar{\theta})$ and $P_\epsilon^{-1}(\tau_c) \cap \chi(\bar{\theta})$ as the *turnstile lobes* of that phase slice. Again we assume that each two-dimensional turnstile lobe and its image is entirely inside or outside the core, with the understanding that we can redefine the lobes as in the time-periodic case if the need arises. Hence the turnstile lobes of each phase slice are either entrained into or detrained from the core under P_ϵ ; again we stress these lobes are the *only* mechanism for entrainment and detrainment relative to the core sequence. Previous investigators have studied escape from bounded to unbounded motion in low-dimensional chaotic systems by focusing on such notions as overlap of neighbouring resonance zones (see for example Goggin and Milonni 1988a, b) and cantori as partial barriers to transport (see MacKay and Meiss 1988). These methods provide only an approximate framework for escape to unbounded motion, in contrast to a theory dealing with the lobes of the homoclinic/heteroclinic tangles, which provides a framework for exact computation of escape from the core, and which, as we shall see, can offer simple explanations of transport phenomena that heretofore have had insufficient explanation. The behaviour of the turnstile lobes motivates the following definition of sets of two-dimensional lobes in a given phase slice, and hence invariant $(l+1)$ -dimensional lobes in $\Sigma^{\theta_{1_0}}$.

Definition 3.5. For $\bar{\theta} \in \mathcal{X}^{l-1}$, $m \in \mathbb{Z}$, $m > 0$, $E(m, \bar{\theta})$ and $D(m, \bar{\theta})$ denote, respectively, the set of lobes in $\chi(\bar{\theta})$ that are entrained into/detrained from the core upon

the m th iterate of P_ε ; for $m < 0$ they denote the set of lobes in $\chi(\bar{\theta})$ that are detrained from/entrained into the core upon the $(|m| + 1)$ th iterate of P_ε^{-1} . We refer to $E(m, \bar{\theta})$ and $D(m, \bar{\theta})$ as sets of *entraining* and *detraining* lobes, respectively. Invariant $(l + 1)$ -dimensional lobes in Σ^{θ_0} are defined by:

$$L(m) = \{L(m, \bar{\theta}) \mid \bar{\theta} \in \mathcal{X}^{l-1}\}$$

where $L = E$ or D .

From previous remarks, $E(m, \bar{\theta}) \cup D(m, \bar{\theta})$ denotes the set of lobes in $\chi(\bar{\theta})$ between $P_\varepsilon^{-(m-1)}(\tau_c) \cap \chi(\bar{\theta})$ and $P_\varepsilon^{-m}(\tau_c) \cap \chi(\bar{\theta})$. In the multiple frequency case the number of lobes in $E(m, \bar{\theta})$ and $D(m, \bar{\theta})$, denoted $\mathcal{N}_E(m, \bar{\theta})$ and $\mathcal{N}_D(m, \bar{\theta})$, respectively, in general depend on $m, \bar{\theta}$, and whether the set is entraining or detraining. At times we will want to refer to individual lobes, which we denote by $E(m, \bar{\theta}; i)$, $D(m, \bar{\theta}; i)$. From previous comments on the dynamics:

$$\mathcal{N}_L(m, \bar{\theta}) = \mathcal{N}_L\left(1, \bar{\theta} + 2\pi \frac{\omega}{\omega_l}(m-1)\right). \quad (3.4)$$

Hence, determining the number of turnstile lobes in each phase slice determines the number of lobes in each set of lobes for each phase slice; further it tells us how many lobes in each phase slice are entrained into/detrained from the core under P_ε . This number can be determined visually in each phase slice for the examples in figure 10 by the portrayals of τ_c and $P_\varepsilon^{-1}(\tau_c)$; however, one need not determine these quantities from plots of PIMS. In the near-integrable case, the Melnikov function provides a computational tool for determining these quantities. If $\mathcal{N}(\bar{\theta})$ is the number of zeros of the Melnikov function in $\chi(\bar{\theta})$ between (and not counting) τ_c and $P_\varepsilon^{-1}(\tau_c)$, then

$$\mathcal{N}_E(\bar{\theta}) + \mathcal{N}_D(\bar{\theta}) = \mathcal{N}(\bar{\theta}) + 1$$

where for simplicity of notation we have written $\mathcal{N}_E(1, \bar{\theta}) \equiv \mathcal{N}_E(\bar{\theta})$ and $\mathcal{N}_D(1, \bar{\theta}) \equiv \mathcal{N}_D(\bar{\theta})$. If $\mathcal{N}(\bar{\theta}) + 1$ is even then $\mathcal{N}_E(\bar{\theta}) = \mathcal{N}_D(\bar{\theta}) = (\mathcal{N}(\bar{\theta}) + 1)/2$. If $\mathcal{N}(\bar{\theta}) + 1$ is odd then one of $\mathcal{N}_E(\bar{\theta})$, $\mathcal{N}_D(\bar{\theta})$ equals $\mathcal{N}(\bar{\theta})/2$, the other $\mathcal{N}(\bar{\theta})/2 + 1$; which one equals which can be determined by the sign of the Melnikov function between the zeros. If $s_c(\bar{\theta})$ denotes the s value of $\tau_c \cap \chi(\bar{\theta})$, then $(s_c(\bar{\theta} + 2\pi(\omega/\omega_l)) + (2\pi/\omega_l))$ is the s value of $P_\varepsilon^{-1}(\tau_c) \cap \chi(\bar{\theta})$, and hence one determines $\mathcal{N}(\bar{\theta})$ by computing the number of zeros of $M(s, \bar{\theta}_1, \dots, \bar{\theta}_{l-1}, \theta_0; \nu)$ between $s_c(\bar{\theta})$ and $(s_c(\bar{\theta} + 2\pi(\omega/\omega_l)) + (2\pi/\omega_l))$. When systems are not near-integrable, one can resort to explicit simulation of two-dimensional lobe structures in a sequence of time slices. We discuss and present computer simulations of lobe structures in section 3.3. Once one chooses a $\text{PIP} \in \tau_c$ for the n th time slice, $\tau_c \cap \chi(\theta_0 + 2\pi(\omega/\omega_l)n)$, one easily determines the $\text{PIP} \in P_\varepsilon^{-1}(\tau_c)$ by applying P_ε^{-1} to the $\text{PIP} \in \tau_c$ in the $(n + 1)$ th time slice, $\tau_c \cap \chi(\theta_0 + 2\pi(\omega/\omega_l)(n + 1))$. Note that the above Melnikov prescription ignores the non-generic case where any of the zeros of M are also local extrema of M , corresponding to a nearby (in parameter space) occurrence of manifolds which touch but do not cross: here numerical simulation is needed to determine whether the manifolds really cross at the desired parameters.

The sets of lobes map according to

$$P_\varepsilon(L(m, \bar{\theta})) = L\left(m-1, \bar{\theta} + 2\pi \frac{\omega}{\omega_l}\right). \quad (3.5)$$

The action of P_ε on m follows by the definition of the lobes, and is hence not very informative; this is because we have chosen a lobe labelling that incorporates the dynamics, which will facilitate later discussions of transport. Note though that a determination of the number of turnstile lobes in each phase slice in principle specifies the number of lobes in each set of lobes, and thus along with equation (3.5) specifies how individual lobes map from one to another; in particular, it specifies how many lobes are entrained and detrained with each iterate of P_ε . For all useful transport expressions, however, we will be interested in how *sets* of lobes map from one to another.

Projecting onto X immediately labels individual lobes and sets of lobes in X for any time sample:

$$e(m, n; i) = \hat{X} \left(E \left(m, \theta_0 + 2\pi \frac{\omega}{\omega_l} n; i \right) \right) \quad (3.6)$$

and similarly for individual detraining lobes and sets of lobes (see figure 13). The projection further specifies how sets of lobes map from one to another under the sequence of maps $\{T_\varepsilon(\cdot; n) \mid n \in \mathbb{Z}\}$:

$$T_\varepsilon(L(m, n); n) = L(m - 1, n + 1) \quad (3.7)$$

where now $L = e$ or d . Note that, for a general bi-infinite sequence of non-repeating maps, to specify the number of turnstile lobes at each time sample necessitates determining a countable infinity of numbers. However, multiple frequency perturbations offer a fundamental simplification due to the compactness of T^{l-1} : if one determines the number of turnstile lobes in each phase slice $\chi(\bar{\theta})$ for all $\bar{\theta} \in \mathcal{X}^{l-1}$, then by (3.6) this determines the number of turnstile lobes for each map in the non-repeating sequence.

Figure 10 illustrates our description of transport. The defined τ_c and a few iterates of $P_\varepsilon^{\pm 1}$ on τ_c are portrayed; the number of turnstile lobes for each phase slice is immediately deduced by looking at the number of pips between τ_c and $P_\varepsilon^{-1}(\tau_c)$ in the phase slice. The $\omega_1 = \omega_2 \equiv \omega$ ($A_1 = A_2 = 1$ until otherwise stated) case in figure 10(a) shows a single pair of entraining/detraining turnstile lobes for all $\theta_1 \in [0, 2\pi)$ (except for the isolated value $\theta_1 = \pi$), and the sets of lobes map according to

$$P_\varepsilon(L(m, \theta_1)) = L(m - 1, \theta_1 + 2\pi) = L(m - 1, \theta_1) \quad \forall \theta_1 \in [0, 2\pi)$$

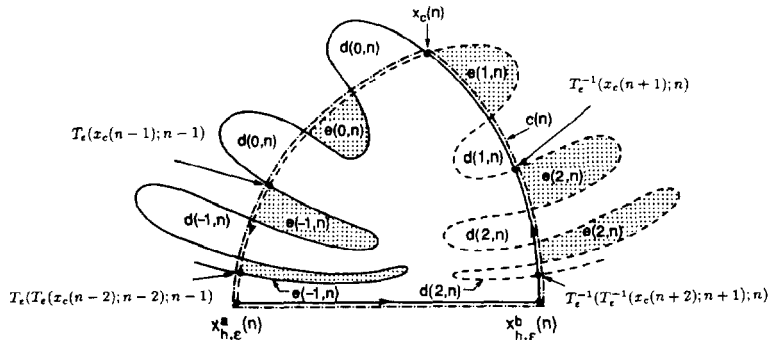


Figure 13. Portraying the core and labelling the sets of lobes in X for the n th time sample of the two frequency ovf flow ($x_c(n) = \hat{X}(\tau_c \cap \chi(\theta_{10} + 2\pi(\omega_1/\omega_2)n))$, $x_{h,e}^i(n) = \hat{X}(\tau_\varepsilon^i \cap \chi(\theta_{10} + 2\pi(\omega_1/\omega_2)n))$ for $i = a, b$, and $T_\varepsilon^{-1}(\cdot; n)$ maps from $t = (n + 1)(2\pi/\omega_2)$ to $t = n(2\pi/\omega_2)$).

where $L = E$ or D . One can easily visualize the lobes mapping under P_ϵ in figure 10(a) by recognizing that $P_\epsilon^{-1}(\tau_c)$ maps to τ_c ; the mapping is portrayed in all of figures 10 (a)–(g) by showing successive images of a single two-dimensional lobe, represented in the (s, θ_1) plane by a dot between the lobe's two bounding PIMs. The two-dimensional lobes of figure 10(a) remain in the same phase slice and thus map within an invariant two-dimensional lobe structure. Hence for $\theta_{10} = 0$ we recover the time-periodic result given earlier for the ovf flow. Exact cancellation of two sine functions out of phase by π gives the vertical zero set at $\theta_1 = \pi$.

The $\omega_1 = 2\omega_2 \equiv 2\omega$ case of figure 10(b) has two pairs of entraining/detraining turnstile lobes for all θ_1 (since we are sampling at the slower frequency), except for the isolated values where the PIMs intersect at $\theta_1 = \pi/2, 3\pi/2$ (we henceforth do not keep pointing out the isolated values). The sets of lobes map according to:

$$P_\epsilon(L(m, \theta_1)) = L(m-1, \theta_1 + 2\pi \cdot 2) = L(m-1, \theta_1) \quad \forall \theta_1 \in [0, 2\pi).$$

As in the previous figure, P_ϵ acts on the indice m , but leaves (θ_1, θ_2) unchanged, so that the two-dimensional lobes remain in the same phase slice. This is an example of a general property of the commensurate frequency case: if $\omega_1/\omega_2 = j/k, j \in \mathbb{Z}, k \in \mathbb{Z}$, then there exists an $N \in \mathbb{Z}$ (for example $N = j \cdot k$) such that P_ϵ^N leaves (θ_1, θ_2) unchanged. This operator P_ϵ^N thus acts on the two-dimensional lobes the same way as does the familiar Poincaré map derived from time-periodic vector fields. The price of this correspondence is that the time between iterates of P_ϵ^N , $\Delta t = N2\pi/\omega_2$, can be extremely long. The $2\omega_1 = \omega_2 \equiv 2\omega$ case of figure 10(c) is the same problem as figure 10(b) sampled at the faster frequency (we purposely consider both samplings for comparison). Now there is only one pair of entraining/detraining turnstile lobes for all $\theta_1 \in [0, 2\pi)$, and the transport equations are given by

$$P_\epsilon(L(m, \theta_1)) = L(m-1, \theta_1 + \pi) \quad \forall \theta_1 \in [0, 2\pi).$$

The lobes now map back and forth between two phase slices separated by $\Delta\theta_1 = \pi$. In general for an arbitrary sampling of a commensurate frequency case, the lobes will visit a finite number of θ_1 values in $[0, 2\pi)$.

Figure 10(d) shows an incommensurate frequency case $\omega_1 = g\omega_2 \equiv \omega$, where g is the golden mean, so the sequence of non-autonomous maps is now truly non-repeating. The PIMs appear qualitatively quite similar to the previous commensurate example. In fact, for all θ_1 values there is one pair of entraining and detraining turnstile lobes; hence a non-repeating map sequence can have underlying PIMs with such simple geometry that the lobes map from one to another in a very simple manner. The lobe sets map according to

$$P_\epsilon(L(m, \theta_1)) = L(m-1, \theta_1 + 2\pi g) \quad \forall \theta_1 \in [0, 2\pi).$$

The essential difference in transport compared with the previous example is that the time slices will now visit $\theta_1 \in [0, 2\pi)$ densely, as will always be the case with incommensurate frequencies. Figure 10(e) shows the case $g\omega_1 = \omega_2 \equiv \omega$, where the number of turnstile lobes depends on θ_1 , as can easily be seen by comparing τ_c and $P_\epsilon^{-1}(\tau_c)$ (see Beigie et al (1991b) for explicit expressions of turnstile lobe numbers). The lobe sets map according to

$$P_\epsilon(L(m, \theta_1)) = L(m-1, \theta_1 + 2\pi/g) \quad \forall \theta_1 \in [0, 2\pi).$$

The last two examples deal with simpler and yet generic examples of unequal Melnikov function amplitudes, and hence PIMs that are non-intersecting tori or

spirals. Figure 10(f) shows a large amplitude, slow frequency term modulated by a small amplitude, fast frequency term ($A_1 = (2\pi A_2) = 1$, $\omega_1 = (\omega_2/2\pi) \equiv \omega$). The lobe structure is essentially determined by the large amplitude component, and we sample here at the faster frequency, so only a small set of θ_1 values actually have turnstile lobes (again as easily seen by comparing τ_c and $P_\epsilon^{-1}(\tau_c)$). The lobe sets map according to

$$P_\epsilon(L(m, \theta_1)) = L(m-1, \theta_1+1) \quad \forall \theta_1 \in [0, 2\pi).$$

Figure 10(g) is the previous system sampled at the slower frequency (i.e. $(2\pi A_1) = A_2 = 1$, $\omega_1/2\pi = \omega_2 \equiv \omega$). Now the PIMS are 1-tori, rather than segments of spirals, and there is again a single pair of entraining/detraining turnstile lobes for all $\theta_1 \in [0, 2\pi)$. The lobe sets map according to:

$$P_\epsilon(L(m, \theta_1)) = L(m-1, \theta_1 + (2\pi)^2) \quad \forall \theta_1 \in [0, 2\pi).$$

For all the examples $\mathcal{N}_E(\bar{\theta}_1)$, $\mathcal{N}_D(\bar{\theta}_1)$ are easily determined for all $\bar{\theta}_1 \in [0, 2\pi)$, and this directly determines $\mathcal{N}_E(\theta_{1_0} + 2\pi(\omega_1/\omega_2)n)$, $\mathcal{N}_D(\theta_{1_0} + 2\pi(\omega_1/\omega_2)n)$, and hence $\mathcal{N}_e(n)$, $\mathcal{N}_d(n)$, for all $n \in \mathbb{Z}$. Hence the number of entraining and detraining turnstile lobes is determined for *all* time samples, due to the compactness of \mathcal{X}^{t-1} .

In summary then, the transport in the tangle region of the non-autonomous system (2.2) is understood in terms of the two-dimensional lobes of the n th time slice of Σ^{θ_0} mapping under P_ϵ to the lobes of the $(n+1)$ th time slice, and we now know how to describe such a map.

3.3. Computer simulation of lobe boundaries and qualitative discussion of lobe transport

As in the time-periodic case, there is for quasiperiodic vector fields a generic picture of transport associated with lobes in phase space mapping from one to another in the manner just described. Some essential features of the picture can be described in a precise way, and in the next two sections we will address some of these features. A good prelude to such a study is the computer simulation of the lobe structure. This provides a direct portrayal of transport in the tangle region, and it will make obvious some features of transport that will be studied in later sections. Our sampling method used for computer simulation, which we refer to as a *double phase slice method* (see Beigie *et al* 1991b), contrasts with a previous suggestion by Moon and Holmes (1985) for analysing the dynamics under quasiperiodic vector fields, a double Poincaré map method which essentially attempts to treat the system as periodic. These authors sample an equation like (2.2) (for the two frequency case) only when both $\theta_2 = \bar{\theta}_2$ and $\theta_1 \in [\bar{\theta}_1 - \beta, \bar{\theta}_1 + \beta]$ (for some choice of $\bar{\theta}_1$, $\bar{\theta}_2$ and $\beta \ll 2\pi$), and the results are summed. The time between samples can be much longer with this approach, and there is a ‘fuzziness’ of the resulting structure due to the finite width of the sampling window, 2β . Our approach uses the fact that points map from phase slice to phase slice and simulates stable and unstable manifolds in a given phase slice by initializing local stable and unstable manifolds in two *different* phase slices. *The results are exact representations of lobes, which thus provide the framework for all computational results, such as lobe area and transport (to be discussed later).*

We present a simulation of the two frequency ovp flow with a $1:g^{-1}$ frequency ratio in the forcing term (the oscillating strain-rate field):

$$\psi_{\text{forcing}} = 0.12x_1x_2\{2 \times 0.4 \sin(2t + \theta_{1_0}) + 2g^{-1} \times 0.95 \sin(2g^{-1}t)\} \quad (3.8)$$

where $\theta_{1_0} = 2\pi[8g - 4] \approx 5.933$ (ψ_{forcing} is the streamfunction of the forcing term in the *non-dimensional equations*). Figure 14 shows a sequence of four time slices and how the lobes of fluid map within this sequence. The time-dependent lobe structure in X for any time $t = (2\pi/\omega_2)n$ has many similarities with the invariant two-dimensional lobe structure of the time-periodic case. There is a reasonable looking core region and a bi-infinite sequence of lobes that extend from the inside and outside of the core boundary, each of which gets thinner and longer as its base (i.e.

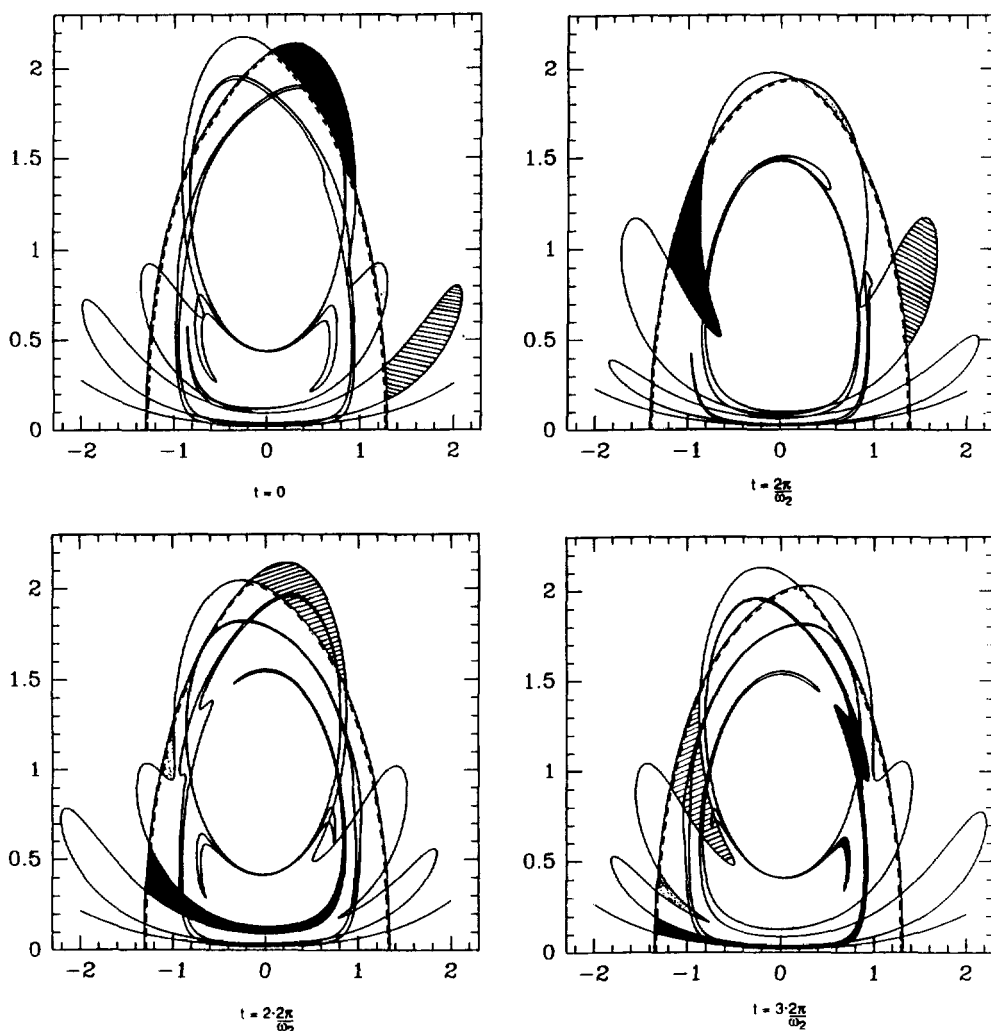


Figure 14. A sequence of four time samples of the ovp flow forced according to equation (3.8). Three material lobes (entraining lobes) are shaded so that we can monitor their transport. The core boundaries are marked by broken lines.

where it extends from the core) gets closer to the points that lie on the normally hyperbolic invariant 1-tori. The thinner and longer the entraining/detraining lobes that extend from inside the core boundary are, the more they wind about the core and wrap around other entraining/detraining lobes and intersect other detraining/entraining lobes (this happens since $W^s(\tau_e^a)$, $W^u(\tau_e^b)$ cannot intersect themselves, but can intersect each other). That the time-dependent lobe structure in X has the same geometrical constraints as the invariant lobe structure of the time-periodic case proves instrumental in easily extending lobe intersection analysis, as discussed in section 4. A notable difference from the single frequency ovf flow, however, is the variation of lobe areas and, as we shall see, this variation can have a significant effect on transport, for it gives one the freedom to enhance or diminish features of transport over finite time scales.

A sequence of snapshots of the lobe structure in X shows notable differences from the time-periodic case. Of course, the lobe structure now varies with each sample time $t = (2\pi/\omega_2)n$ (in a non-repeating manner for incommensurate systems and a repeating manner for commensurate ones). The core region varies in shape and area with each sample, as do lobe areas relative to their ordering with respect to the $\text{PIP} \in \tau_c$ (which does not contradict that lobes of fluid conserve area as they map from one lobe structure to the next). As should be clear from the previous sections, the key to understanding the time-dependent structure is to recognize that it is the intersection of a time slice with an invariant structure in a higher dimensional Poincaré section; for example, we stress how there are no fixed points in the two-dimensional structure, but rather points on a normally hyperbolic invariant 1-torus in Σ^{θ_0} .

As lobes of fluid map in X from one lobe structure to the next, their behaviour is qualitatively similar to that found in the time-periodic case. They stretch in one direction and contract in another to produce the two essential features of lobe transport found in the time-periodic case: entrainment and detraining (back and forth motion across boundaries that were barriers to transport in the unperturbed case) and repeated stretching and folding to produce chaotic advection. These features need to be understood in the framework of the sequence of non-autonomous maps on X , which will be the focus of the next two sections. An essential aspect of the dynamics to recognize right away, however, for both the time-periodic and quasiperiodic case, is the role of the unstable manifold as a dominant structure in chaotic tangles (see Rom-Kedar *et al* (1990) for the time-periodic case): due to the repeated stretching and folding of lobes (see section 5 for the quasiperiodic case), material curves in the tangle region tend to get 'attracted' to the unstable manifold. For Hamiltonian systems there is of course no true attractor; however, though area elements do not shrink, they tend to be stretched in one direction and contracted in another such that the unstable manifold dominates the evolution of material curves. Since $\hat{X}(W^u(\tau_e^b) \cap \chi(\theta_{1_0} + 2\pi(\omega_1/\omega_2)n))$ varies with n , the 'attracting' structure in X is time-dependent. The time-dependent lobe structure is thus the dominant structure by which to understand motion in the tangle region, and it allows one to embrace rather than avoid the time-dependent nature of the more complicated transport issues under quasiperiodic perturbations. We wish to stress the notion of a time-dependent attractor. Often a physical system is a point in an autonomous system phase space, and so there is little motivation to conceptualize anything but the invariant attractor in the autonomous system phase space. In chaotic advection, however, where physical space at sample

times is a time slice of the autonomous system phase space, the time-dependent 'attractor' associated with each time slice does have physical meaning.

4. Entrainment and detrainment

4.1. Instantaneous and average flux

Several features of entrainment and detrainment can be studied via the lobe structures. We start with lobe areas, which provide a measure of *flux* across the core boundaries. For simplicity of discussion, let us first consider Hamiltonian perturbations, i.e. area-preserving map sequences on X .

Definition 4.1.1. The *instantaneous flux* associated with P_e acting on the phase slice $\chi(\bar{\theta})$, denoted by $\phi_E(\bar{\theta})$ and $\phi_D(\bar{\theta})$ for entrainment and detrainment, respectively, is the area of the regions in the phase slice that are entrained into/detrained from the core under P_e , divided by the sampling period.

Hence expressions for instantaneous flux associated with P_e acting on the n th time slice are given by

$$\phi_L\left(\theta_0 + 2\pi \frac{\omega}{\omega_l} n\right) = \frac{\omega_l}{2\pi} \mu\left(L\left(1, \theta_0 + 2\pi \frac{\omega}{\omega_l} n\right)\right) \quad (4.1)$$

where $L = E$ or D , and $\mu(\cdot)$ denotes the area of the set of lobes within the parenthesis. Recall that section 3 explained how to identify the turnstile lobes in each time slice.

Definition 4.1.2. The *instantaneous flux* associated with $T_e(\cdot; n)$ acting on X , denoted by $\phi_e(n)$ and $\phi_d(n)$ for entrainment and detrainment, respectively, is the area of the regions in X that are entrained into/detrained from the core under $T_e(\cdot; n)$, divided by the sampling period.

Hence the expressions for instantaneous flux associated with $T_e(\cdot; n)$ acting on X are given by

$$\phi_L(n) = \frac{\omega_l}{2\pi} \mu(L(1, n)) \quad (4.2)$$

where $L = e$ or d . Since $e(m, n) = \hat{X}(E(m, \theta_0 + 2\pi(\omega/\omega_l)n))$ and $d(m, n) = \hat{X}(D(m, \theta_0 + 2\pi(\omega/\omega_l)n))$, $\phi_e(n) = \phi_E(\theta_0 + 2\pi(\omega/\omega_l)n)$ and $\phi_d(n) = \phi_D(\theta_0 + 2\pi(\omega/\omega_l)n)$, so for the remainder of this section we deal just with slices of the autonomous system phase space, with the understanding of how the flux results apply directly to the non-autonomous system. Note that, in contrast to the single frequency case, instantaneous flux will in general vary from one time slice to the next, and will be different for entrainment and detrainment.

In addition, we will also be concerned with *average flux*.

Definition 4.2. The *average flux* associated with P_e acting on the bi-infinite sequence of time slices, or equivalently with the bi-infinite sequence of maps $\{T_e(\cdot; n) \mid n \in \mathbb{Z}\}$ acting on X , is the average of the instantaneous flux over all visited time slices.

Since P_ε preserves area as it acts on a time slice, the average flux is identical for entrainment and detrainment, and is thus denoted by a single symbol Φ :

$$\Phi(\theta_0) = \frac{1}{2} \left\{ \left\langle \phi_E \left(\theta_0 + 2\pi \frac{\omega}{\omega_l} n \right) \right\rangle_n + \left\langle \phi_D \left(\theta_0 + 2\pi \frac{\omega}{\omega_l} n \right) \right\rangle_n \right\} \quad (4.3)$$

where $\langle \rangle_n$ denotes the average over $n \in \mathbb{Z}$ (when one or more frequencies are mutually commensurate this average will depend on θ_0 , and when the frequencies are all mutually incommensurate it will be independent of θ_0).

For sufficiently small perturbations, the Melnikov approximation of manifold separations in equation (2.7) provides, for any phase slice $\chi(\bar{\theta})$ of Σ^{θ_0} , $\bar{\theta} \in \mathcal{X}^{l-1}$, a first order in ε approximation of areas of the individual turnstile lobes. From the comments in section 2 on uniform approximation of $W^s(\tau_\varepsilon^a)$ and $W^u(\tau_\varepsilon^b)$ on semi-infinite time intervals, for sufficiently small perturbations the boundaries of the turnstile lobes are $O(\varepsilon)$ close to the unperturbed manifold. The area element of one of these lobes, as shown in figure 15, has an area

$$\begin{aligned} dA &= \left[\frac{\varepsilon |M(s, \bar{\theta}_1, \dots, \bar{\theta}_{l-1}, \theta_{l_0}; \nu)|}{\|DH(x_h(-s))\|} + O(\varepsilon^2) \right] [d\lambda(1 + O(\varepsilon))] \\ &= \frac{\varepsilon |M(s, \bar{\theta}_1, \dots, \bar{\theta}_{l-1}, \theta_{l_0}; \nu)|}{\|DH(x_h(-s))\|} d\lambda + O(\varepsilon^2) \end{aligned}$$

where $d\lambda$ is an element of arc length, centred at $x_h(-s)$, along the unperturbed homoclinic or heteroclinic manifold in $\chi(\bar{\theta})$. Using that

$$d\lambda = \frac{ds}{\|DH(x_h(-s))\|}$$

gives

$$\mu(\text{individual turnstile lobe } (\bar{\theta})) = \int_{s_a(\bar{\theta})}^{s_b(\bar{\theta})} |M(s, \bar{\theta}_1, \dots, \bar{\theta}_{l-1}, \theta_{l_0}; \nu)| ds + O(\varepsilon^2) \quad (4.4)$$

where $s_a(\bar{\theta})$ and $s_b(\bar{\theta})$ are the s values of the bounding PIPs of the individual turnstile lobe in $\chi(\bar{\theta})$. Equations (4.1) and (4.4) provide an analytical prescription for determining instantaneous flux, and some calculations are given in Beigie *et al*

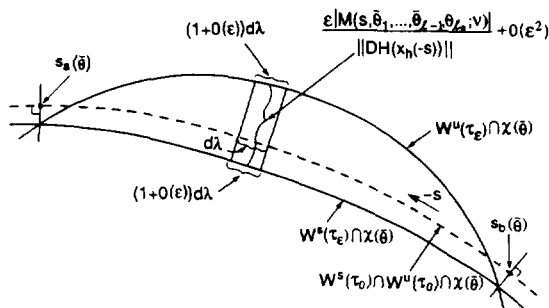


Figure 15. An area element of a two-dimensional turnstile lobe in a given phase slice for an l frequency homoclinic tangle.

(1991b). As derived in Beigie et al (1991b), the average flux is given by

$$\Phi(\theta_0) = \lim_{T \rightarrow \infty} \frac{\varepsilon}{2T} \int_0^T |M(s, \theta_{1_0}, \dots, \theta_{(l-1)_0}, \theta_{l_0}; \nu)| ds + O(\varepsilon^2). \quad (4.5)$$

Expression (4.5) is an integral over a *single* phase slice, and since the Melnikov approximation for manifold separation (2.7) is valid only over a finite time interval, equation (4.5) is not immediately obvious. The expression for $\Phi(\theta_0)$ is really a *sum* over phase slices of the turnstile lobe areas, for which (2.7) is valid for small ε , but this is converted to an integral over a *single* phase slice by elementary periodicity properties of the Melnikov function. Note that for the heteroclinic core there can of course be multiple sets of turnstiles (potentially one set associated with each unperturbed heteroclinic orbit) and the average flux (4.5) applies to only one set of turnstiles.

One can thus use the generalized Melnikov function to perform a detailed study of average flux as a function of system parameters, especially to compare single and multiple frequency forcing. There are three ingredients to understanding such a comparison:

- (1) the relative scaling factors are frequency dependent;
- (2) a sum of sinusoidal functions interfere;
- (3) one must choose a normalization of perturbation amplitudes in order to compare single and multiple frequency systems.

Let us compare single and two frequency forcing in the following way (from our two frequency comparisons, the extension to l frequency comparisons should be clear). Consider the perturbation in equation (2.8) with $F_0(x; \mu) = 0$ (as is the case in the ovp flow). For any frequency pair (ω_1, ω_2) , study the dependence of Φ on F_1 , with F_2 chosen by some normalization condition (for the ovp flow $F_i = \omega_i f_i$, $i = 1, 2$, as should be clear from the appendix). For example, suppose $F_1 + F_2 = 1$. For equal relative scaling factors, both single frequency cases ($F_1 = 0$ and $F_1 = 1$) will correspond to absolute maxima, due to interference effects (see the solid line in figure 16(a)). For unequal relative scaling factors (say $A_1/F_1 > A_2/F_2$ without loss of generality), the single frequency case associated with the larger relative scaling factor (the case $F_1 = 1$) will correspond to an absolute maximum of average flux (see the broken line in figure 16(a)). This holds for arbitrary relative scaling factors (see Beigie et al 1991b). *The interpretation for this normalization is then this: for a two frequency perturbation, in general one of the frequencies has a greater relative scaling factor and hence a larger contribution to manifold separation and lobe area for a given forcing amplitude; to include the other frequency in the perturbation only detracts from average flux—the best one can do is the single frequency case with the larger relative scaling factor.* Of course, the comparisons depend on the choice of normalization, and it is possible to choose a normalization such that the perturbation amplitudes are large enough at mid-range values of $F_1 \in [0, 1]$ that the effects due to interference and different relative scaling factors are outweighed, and then the multiple frequency case has maximal flux. Preferably the dynamical system equations describe a physical situation which motivates a normalization. For example, as shown in Rom-Kedar et al (1990), the ovp flow can be thought of as a vortex pair moving in a wavy-wall channel. Suppose we choose to keep the root-mean-square wall amplitude constant for all $f_i \in [0, 1]$ for any given (ω_1, ω_2) . For the long wavelength limit of the wall amplitude oscillations, the wall amplitudes

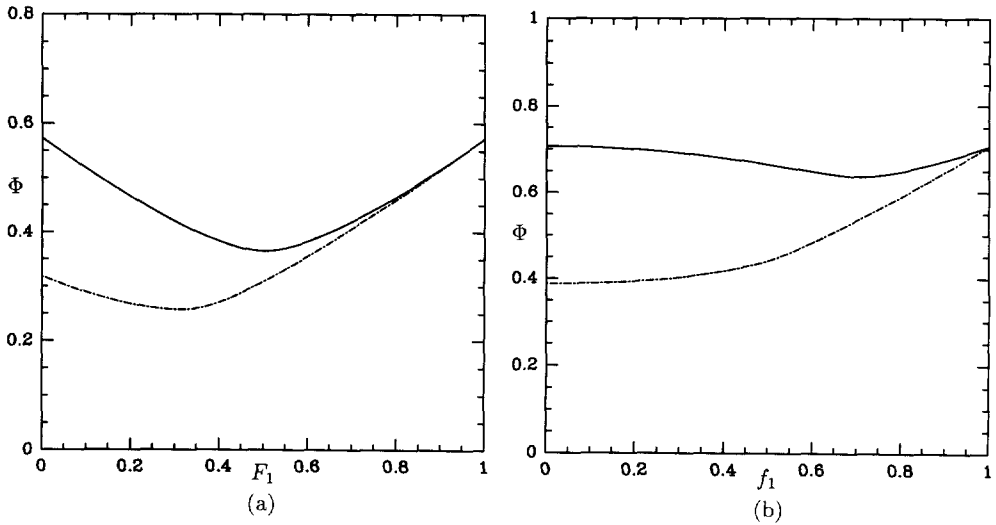


Figure 16. (a) Average flux as a function of F_1 for $F_1 + F_2 = 1$, with $(\omega_1, \omega_2) = (1.48, 0.78)$, $(A_1/F_1, A_2/F_2) = (1.80, 1.80)$ for the solid line and $(\omega_1, \omega_2) = (1.48, 0.66)$, $(A_1/F_1, A_2/F_2) = (1.80, 1.00)$ for the broken line (A_i/F_i is taken from the OVP flow: see the appendix). (b) Average flux for the OVP flow as a function of f_1 for $f_1^2 + f_2^2 = 1$, with $(\omega_1, \omega_2) = (1.00, 1.94)$, $(A_1/f_1, A_2/f_2) = (2.22, 2.22)$ for the solid line and $(\omega_1, \omega_2) = (1.00, 0.75)$, $(A_1/f_1, A_2/f_2) = (2.22, 1.22)$ for the broken line. In both plots the vertical scale is per unit ε and $\theta_{10} = \theta_{20} = 0$.

are simply related to the perturbation amplitudes, and the normalization becomes $f_1^2 + f_2^2 = 1$. Figure 16(b) shows some plots of average flux for two choices of (ω_1, ω_2) . The profiles are similar to those in figure 16(a), except that the interference dips are not quite so pronounced as with the previous normalization ($F_1 + F_2 = 1$). Here too average flux is maximal in a single frequency limit. Note that for perturbations (2.8) the single frequency (ω_i) flux is proportional to A_i , and hence to $F_i \text{RSF}(\omega_i)$, where $\text{RSF}(\omega_i)$ denotes the relative scaling factor. Thus, whereas for the first normalization the maximal single frequency limit corresponds to the frequency ω_i with greater value of $\text{RSF}(\omega_i)$, for the second normalization it corresponds to the frequency ω_i with greater value of $\omega_i \text{RSF}(\omega_i)$.

Perhaps more important than making a comparison between the single and multiple frequency case, however, is the ability to search through the parameter space of the multiple frequency problem. Hence, just as one can use the generalized Melnikov function to study where in parameter space chaos can occur (see Ide and Wiggins (1989), and section 5), one can in addition search these chaotic regions to study the magnitude of the average flux. The amount of material in phase space that is available for repeated stretching and folding is related to the average flux (i.e. to average lobe areas), and so this magnitude can be viewed as one measure of how 'chaotic' a system is. One could of course produce endless plots like those in figure 16 to sweep through parameter space, but it is not our goal here to do a detailed numerical study. Before one performs such a study, however, we cannot overemphasize the usefulness of the relative scaling function. A plot of this function, along with a chosen normalization, immediately indicates the effectiveness of each frequency at producing manifold separations, which allows one to deduce such practical results as which frequencies have the greatest contribution to average flux,

and when one frequency component will overwhelm the other, so that the two frequency forcing is essentially single frequency forcing.

We emphasize that the analytical expressions we have provided are valid through $O(\varepsilon)$, and hence are good approximations only in the near-integrable case. An exact determination of lobe areas, and hence flux, necessitates straightforward, yet laborious, computation. One needs to determine the turnstile lobe boundaries in each time slice, and this can be done by the previously mentioned *double phase slice method*. Then it is a simple matter of computing the areas between boundaries in a plane. Other investigators (for example MacKay and Meiss 1988) point out in the context of time-periodic perturbations that the Melnikov function provides an *exact* expression for flux through the *unperturbed separatrix*. However, as ε grows, the core boundary formed by segments of stable and unstable manifolds can move further away from the unperturbed separatrix, and hence the difference between the flux associated with the unperturbed separatrix and the flux associated with the core sequence can grow (the former becoming smaller and smaller compared with the latter). The larger flux associated with the core sequence is a truer measure of the magnitude of the back and forth motion in the tangle region; further, from previous comments in section 3 it is the perturbed manifolds which should provide the criterion for entrainment and detrainment and act as a 'last frontier' between bounded and unbounded motion.

For non-Hamiltonian perturbations there is the additional complication that the two-dimensional turnstile lobes do not conserve area as they are entrained into or detrained from the core under P_ε . The definitions and analysis for flux extend to non-Hamiltonian perturbations if we replace everywhere the area of the turnstile lobes by the area of the *images* of the turnstile lobes. This replacement introduces *no* changes to the near-integrable $O(\varepsilon)$ approximations for flux since the area of the turnstile lobes and their images are the same through $O(\varepsilon)$. When the system is not near-integrable, one needs to resort to numerical simulation of the boundaries of the *images* of the turnstile lobes.

4.2. Addressing basic transport issues by considering the geometry of lobes and their variation of areas

The average flux provides a measure of the back and forth motion across a reasonable boundary in the tangle region. However, to answer basic questions about transport in phase space necessitates a consideration of lobe geometry and the variation of lobe areas. For example, in the case of a forced system whose unperturbed separatrix divides phase space into bounded and unbounded motion, a fundamental concern (with a variety of physical applications) is the probability of escape from the core after n time samples. This necessitates a consideration of *lobe content*, rather than just lobe areas, which can be determined by addressing the intersection of lobes with each other and with the core. An additional concern (motivated for example by the semiclassical treatment of molecular dissociation) is the probability of escape for points that lie on a particular level set of the unperturbed Hamiltonian. This necessitates a consideration of the intersection of lobes with the level sets of the unperturbed Hamiltonian. Both these transport problems are affected not only by the average of lobe areas, but by the *variation* of lobe areas, and the geometry of lobes.

Let us consider these transport problems. As mentioned in section 3.3, entraining and detraining lobes can intersect each other; from these intersections we obtain a more detailed knowledge of transport in the tangle region. For example, the $t=0$ sample of figure 14, which we present again here in figure 17 with labels on some lobes, shows lobe $d(3,0)$ intersecting with lobe $e(1,0)$; the fluid in $d(3,0) \cap e(1,0)$ will not only be entrained into the core in the next time sample, it will then be detrained from the core between the $n=2$ and $n=3$ time sample. Similarly $e(-2,0)$ intersects $d(1,0)$: the fluid in $e(-2,0) \cap d(1,0)$ will be detrained from the core in the next time sample and was entrained into the core between the $n=-3$ and $n=-2$ time sample. These concerns relate to the *content* of those lobes about to be entrained/detrained in the next time sample: examining the intersections of these lobes with other lobes provides an understanding of the history of the fluid particles, i.e. when they might have been previously entrained/detrained or in the future when they may be entrained/detrained. An example of such a concern is our first question about escape from the core. We begin by examining this question in the context of the multiple frequency ovf flow, i.e. an area-preserving system whose unperturbed solution contains a pair of heteroclinic orbits that divide bounded and unbounded motion. We can phrase the transport problem in the following way: suppose at $t=0$ fluid of one type (A) is in the core, and fluid of another type (B) is outside the core: then at any time $t = (2\pi/\omega_I)n$ how much of fluid A and B is inside and outside the core, i.e. how much of fluid A and B have been *transported* out of and into the core, respectively? For the ovf system, the $e(m,n)$ lobes for $m \geq 1$ and the $d(m,n)$ lobes for $m \leq 0$ are contained entirely outside the core. The remaining lobes will in general lie both inside and outside the core, except for the lobes $e(0,n)$ and $d(1,n)$ by our previous turnstile assumption (see figure 17). The particle content of a lobe of fluid is of course determined by its intersection with the core at time $t=0$. If $\mu_A(\cdot)$, $\mu_B(\cdot)$ denote the areas of the pieces

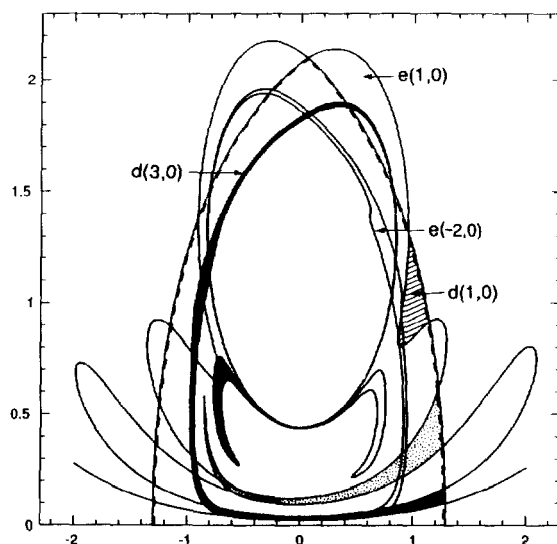


Figure 17. The lobe structure of the $t=0$ sample of figure 14, shown again with the necessary labelling (this time some detraining lobes are shaded).

of the lobe in the parenthesis containing fluid A and B , respectively, then

$$\begin{aligned}\mu_A(e(m, 0)) &= \mu(e(m, 0) \cap c(0)) \\ \mu_B(e(m, 0)) &= \mu(e(m, 0)) - \mu(e(m, 0) \cap c(0)) \\ \mu_A(d(m, 0)) &= \mu(d(m, 0) \cap c(0)) \\ \mu_B(d(m, 0)) &= \mu(d(m, 0)) - \mu(d(m, 0) \cap c(0)).\end{aligned}\tag{4.6}$$

These expressions can be rewritten in terms of lobe intersections alone. For $m \geq 1$:

$$\begin{aligned}\mu_A(e(m, 0)) &= 0 \\ \mu_B(e(m, 0)) &= \mu(e(m, 0)) \\ \mu_A(d(m, 0)) &= \mu(d(m, 0)) - \sum_{i=1}^{m-1} \mu(d(m, 0) \cap e(i, 0)) \\ \mu_B(d(m, 0)) &= \sum_{i=1}^{m-1} \mu(d(m, 0) \cap e(i, 0))\end{aligned}\tag{4.7}$$

where for $m = 1$ we replace the two sums by zero. The last two equations follow from the fact that the stable manifold cannot intersect itself, so the only way for $d(m, 0)$ ($m \geq 1$) to be outside the core is for it to be in the $e(i, 0)$ lobes for $i \geq 1$ (see figures 13 or 17). The $i \geq m$ intersections are disallowed because of our assumption that the detraining turnstile lobes are completely inside the core (which is valid for the ovp flow up to quite large perturbations).

Now we wish to specify the change of core content between any two time samples. If $\Delta A_{\text{core}}(n)$ and $\Delta B_{\text{core}}(n)$ are the area changes under $T_\epsilon(\cdot; n)$ of fluid type A and B , respectively, in the core, then

$$\begin{aligned}\Delta A_{\text{core}}(n) &= \mu_A(e(1, n)) - \mu_A(d(1, n)) \\ \Delta B_{\text{core}}(n) &= \mu_B(e(1, n)) - \mu_B(d(1, n)).\end{aligned}\tag{4.8}$$

From previous comments, for all positive $i \in \mathbb{Z}$ and all $n \in \mathbb{Z}$

$$\begin{aligned}\mu_A(e(i, n)) &= 0 \\ \mu_B(e(i, n)) &= \mu(e(i, n)).\end{aligned}\tag{4.9}$$

To deal with $d(i, n)$, recall that $T_\epsilon(\cdot; n)$ is area-preserving for all $n \in \mathbb{Z}$, so that

$$\begin{aligned}\mu_A(\cdot) &= \mu_A(T_\epsilon(\cdot; n)) \\ \mu_B(\cdot) &= \mu_B(T_\epsilon(\cdot; n)).\end{aligned}\tag{4.10}$$

Using (4.10) and (4.6) gives expressions for the $d(i, n)$ lobes which, when plugged into (4.8) along with (4.9), give

$$\begin{aligned}\Delta A_{\text{core}}(n) &= -\mu(d(1+n, 0) \cap c(0)) \\ \Delta B_{\text{core}}(n) &= \mu(e(1, n)) - \mu(d(1, n)) + \mu(d(1+n, 0) \cap c(0)).\end{aligned}\tag{4.11}$$

This specifies the area changes in terms of areas of lobes and of intersections of lobes with the core. Alternatively, one could use (4.7) instead of (4.6) to obtain

$$\begin{aligned}\Delta A_{\text{core}}(n) &= -\mu(d(1, n)) + \sum_{j=0}^{n-1} \mu(d(1, n) \cap e(-j, n)) \\ \Delta B_{\text{core}}(n) &= \mu(e(1, n)) - \sum_{j=0}^{n-1} \mu(d(1, n) \cap e(-j, n))\end{aligned}\tag{4.12}$$

(for $n = 0$ replace the sums over j by zero). This specifies the area changes in terms of areas of turnstile lobes and of intersections of turnstile lobes with other lobes. *The right-hand side of (4.12) has the same form as in the time-periodic case with the second lobe argument n simply going along for the ride. This is due to the fact that, as mentioned in section 3, the geometrical constraints of the time-dependent lobe structures in X for the quasiperiodic case are the same as those of the invariant lobe structure of the time-periodic case.* Summing the first expression of (4.11) or (4.12) over the first N time samples and dividing by the initial core area gives the percentage of material initially in the core that escapes after N time samples.

It is particularly easy to discuss lobe content in the ovp example since there is no re-entrainment (once fluid escapes the core, it never re-enters) and the map sequence is area-preserving. For Hamiltonian systems which involve re-entrainment, to specify area changes in terms of intersections of lobes with other lobes is more difficult. However, to specify area changes in terms of intersections of lobes with the core is not more difficult: one still maps lobes back to the $t = 0$ sample and sees how much of the lobe is inside and outside the $t = 0$ core. Since the lobes are no more fundamental an entity than the core, to specify area changes in terms of intersections of lobes with the core is no less fundamental than specification in terms of intersections of lobes with other lobes. Further, if one wishes to compute a lobe's content, it is easier to deal with a single intersection of the lobe mapped back to $t = 0$ with the core at $t = 0$ than with expressions involving a sum of intersections of the lobe with other lobes. Nevertheless, there is still interest in understanding the geometry of intersections of lobes, and we discuss in Beigie *et al* (1991b) how the general transport formalism of Rom-Kedar and Wiggins (1990) in terms of lobe intersections (for both area-preserving and non-area-preserving maps) extends easily to map sequences.

The main difference between single and multiple frequency systems in regard to this transport problem is the *variation* of lobe areas: *given a fixed average flux, the variation of lobe areas gives one the freedom to enhance or diminish transport in a given direction over a finite time.* For example, if one wishes to enhance immediate detraining over the first few time samples, one can choose a perturbation time dependence such that the first few detraining lobes are large (relative to the mean) and the first few entraining lobes are small (which forces the first few detraining lobes to start off mainly in the core and hence take more of the initial core material with it). Of course the fact that average flux is typically worse in multiple frequency systems than in a corresponding single frequency system (as explained earlier) tends to diminish the transport in both directions, but over a finite time interval this can be outweighed by the above effect due to variation of lobe areas.

The second transport problem we introduced in this section deals with the probability of escape for points that lie on a *particular level set of the unperturbed Hamiltonian*. For example, Noid and Stine (1979) and Goggin and Milonni (1988a, b) study dissociation of diatomic molecules due to forcing with one and two lasers, as described by the Morse oscillator; a semiclassical approach to this problem is concerned with the probability that an ensemble of points at a particular level set of the unperturbed Hamiltonian (for example the one corresponding to the quantum mechanical ground state) escape from bounded to unbounded motion. They find in their numerical investigations that the two laser system tends to *enhance* dissociation (i.e. escape from a particular level set to unbounded motion). This contrasts with our earlier result that two frequency systems tend to diminish the average flux,

which one might think would *hinder* dissociation. However, average flux is only one of many factors in this transport problem; one must also consider the geometry of lobes and the variation of lobe areas. For example, though the absolute value of the two frequency Melnikov function may have a lower average over s than in a corresponding single frequency case, for some normalizations it can have a larger maximum, and hence the lobes may better penetrate inner level sets of the unperturbed Hamiltonian. Further, as mentioned briefly in Noid and Stine (1979) and elaborated upon by Goggin and Milonni (1988b), additional forcing frequencies entail additional resonance bands within the core. Though no mention is made of the homoclinic tangle, it is clear that more efficient destruction of KAM tori within the core can allow the lobes of the homoclinic tangle to better penetrate inner level sets. Additionally, for penetration of a fixed unperturbed level set, the previously discussed pulsation and ‘breathing’ of the lobe structure can be significant. We see then that though average flux may decrease, there are other factors that can affect the ability of the lobes to penetrate a particular inner level set in a multiple frequency problem. As a simple example, consider figure 18, which shows again in (a) some lobes for the $n = 2$ time slice of the system portrayed in figure 14, and in (b) some lobes of a corresponding single frequency system ($f_1^2 + f_2^2$ is the same for both systems). Though the two frequency system has a smaller average lobe area, its first detraining lobe is large relative to the mean and intersects with more inner level sets than does the first detraining lobe of the single frequency case. Of course, this does not imply that the first N detraining lobes of this two frequency system intersect with more of the level set shown than do the lobes of the single frequency system; we are focusing for visual simplicity on the first detraining lobe. We should point out that the studies of Noid and Stine (1979) and Goggin and Milonni (1988b) consider a fairly specific situation: they focus on the *threshold intensity* needed for the ground state level set to achieve any dissociation in the long time limit. Escape over *all* time scales (including small and medium time scales) for a forcing amplitude that does not happen to be right at threshold for the concerned level set is a more complicated problem for which conclusions are less easily drawn and a study in terms of lobe dynamics is motivated. In particular, one would expect the variation of lobe areas to affect any studies on finite time scales.

The framework in terms of lobe dynamics is easily written down. One can write for the ovp flow (or equivalently the Morse oscillator, or any system with similar geometry) an expression for probability of escape for this second transport problem

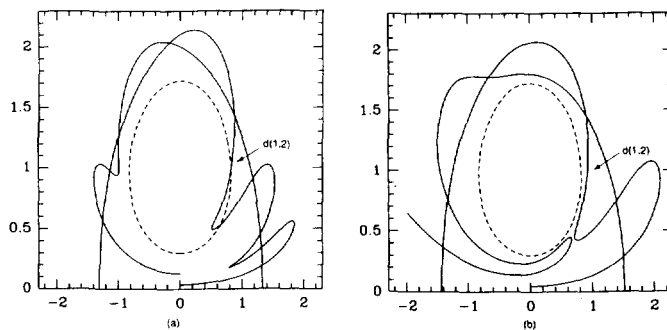


Figure 18. Some lobes in the $n = 2$ time sample of (a) the system portrayed in figure 14 and (b) a corresponding single frequency system ($\omega = 2$).

that is quite similar to the first expression of equation (4.11). If \mathcal{L}_0 denotes the level set of interest of the unperturbed Hamiltonian, and $\Delta\mathcal{L}_0(n)$ denotes the number of points initially on the level set that escape under $T_\varepsilon(\cdot; n)$, where the points are initialized on the level set according to some density, then

$$\Delta\mathcal{L}_0(n) = \mu(d(1+n, 0) \cap \mathcal{L}_0) \quad (4.13)$$

where $\mu(\cdot)$ denotes the number of points on the curve within the parenthesis. Summing $\Delta\mathcal{L}_0(n)$ over the first \mathcal{N} time samples and dividing by $\mu(\mathcal{L}_0)$ gives the percentage of points on the level set that have escaped after \mathcal{N} time samples. Equations (4.11) to (4.13) thus provide a framework for computing answers to the two transport problems. *Once one recognizes the mechanism for, and topology of, transport, there are rich possibilities for determining answers to transport questions by explicit simulation of a finite number of lobes in the homoclinic/heteroclinic tangle.* We conclude this section by remarking that a more detailed study of these transport problems, especially the second one, is given in Beigie and Wiggins (1991) in the context of a Morse oscillator.

5. Some remarks on the chaotic nature of the dynamics

For quasiperiodic systems (2.2), the map of X from time $t = (2\pi/\omega_l)n$ to time $t = (2\pi/\omega_l)(n+1)$, $T_\varepsilon(\cdot; n)$, depends on n . So we of course cannot develop the usual two-dimensional horseshoe map construction. *So in what sense is the dynamics in X chaotic?* One can imagine rather than a single horseshoe map, a bi-infinite sequence S_H of different 'horseshoe maps' $H(\cdot; j): X \rightarrow X$,

$$S_H = \{\dots, H(\cdot; -j), \dots, H(\cdot; -1), H(\cdot; 0), H(\cdot; 1), \dots, H(\cdot; j), \dots\}$$

and a bi-infinite sequence S_D of different domains $D(j) \in X$,

$$S_D = \{\dots, D(-j), \dots, D(-1), D(0), D(1), \dots, D(j), \dots\}$$

such that $H(D(j); j)$ intersects $D(j+1)$ in the shape of a horseshoe (see figure 19). There is thus a sequence of formed horseshoes landing on different regions of X : each time the horseshoe happens to land on the region that will next form a horseshoe, and land in such a way that the stretched direction 'aligns' with the direction about to be stretched. We refer to this as a *travelling horseshoe map sequence*. It is clear that this map sequence retains the essential ingredient of chaos—repeated stretching and folding, and hence sensitive dependence on initial conditions. Though our discussion here is heuristic, figure 19 should make apparent what we mean. As done in Beigie *et al* (1991b), a rigorous construction of a travelling horseshoe map sequence can be made for systems of the form (2.2) which

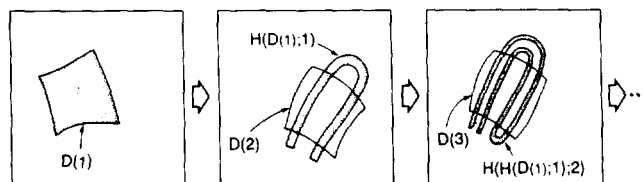


Figure 19. A travelling horseshoe map sequence.

possess an invariant lobe structure. What essentially happens is that the $(l+1)$ -dimensional lobes repeatedly stretch and fold in a violent manner *normal* to the normally hyperbolic $(l-1)$ -tori; hence one can establish the existence of some $(l+1)$ -dimensional domain whose image, after a sufficient number of iterates of P_ϵ , intersects the original domain in the shape of a horseshoe in each phase slice $\chi(\bar{\theta})$, $\bar{\theta} \in \mathcal{X}^{l-1}$ (see Wiggins 1988). A sequence of time slices thus defines a sequence of two-dimensional domains that satisfy the travelling horseshoe map criterion. The chaotic dynamics is thus understood relative to a *time-dependent* Cantor set of points (described rigorously in Beigie et al (1991b)).

We reiterate our previous comments on searching parameter space to determine when one can and cannot have chaos. From previous remarks, such a search is equivalent to a search in parameter space to determine whether a lobe structure exists, which in the near-integrable case is equivalent to determining whether zero sets of the Melnikov function exist. We stress that such a search can have simple, practical consequences. For example, in commensurate two frequency systems the lobe structures, and hence Cantor sets, can have gaps in θ_1 (refer back to figure 4(h)): hence, for a given sampling phase θ_{2_0} , the choice of θ_{1_0} affects whether the system has chaotic dynamics (if θ_{1_0} lies in a gap there will be no chaos). *Thus a shift in the relative initial phase $\theta_{1_0} - \theta_{2_0}$ can suppress chaos.*

In the introduction we remarked that the horseshoe map construction for time-periodic vector fields (or equivalently the travelling horseshoe map construction for quasiperiodic vector fields) is the simplest framework for understanding enhanced stretching in chaotic tangles, and that a study of lobe behaviour provides a more detailed understanding of stretching, one which includes an appreciation of the topology of stretching in chaotic tangles. Adequate discussion of this issue requires substantial additional set-up, so we merely refer the reader to the time-periodic study in Beigie et al (1991a). The quasiperiodic extension of this study should be fairly clear from the present discussion.

6. Conclusions and explanation of extending the analysis to more general time dependences

On the one hand, as is the case with turbulent fluid flows, the behaviour of nonlinear physical systems is often so complex that one can obtain only a very qualitative or partial understanding of the motion. On the other hand, there do exist nonlinear systems which display complicated dynamics, and yet with which one can obtain a fundamental and detailed global picture of the dynamics. At present, there is a wide gap between these two types of systems, and one would of course like to bridge this gap by extending some of the initial paradigms of low-dimensional dynamical systems analysis to more complicated scenarios. It is in this spirit that we have extended some notions of transport in phase space associated with the classical Poincaré map reduction of two-dimensional periodically forced systems to apply to a bi-infinite sequence of non-autonomous maps derived from a quasiperiodic two-dimensional vector field. In a certain sense, this is a mild extension, since it still deals with a highly limited class of perturbation time dependences. We have deliberately focused on a simple extension (especially the two frequency case) since any non-trivial advance past the ubiquitous time-periodic case immediately entails a fundamental departure in the analysis (going from maps to sequences of maps) and

the new concepts in the simple two frequency extension are for the most part robust (as we shall describe momentarily), applicable to more general time dependences. Our goal then has been to discuss these new concepts in the context of a simple class of systems.

Suppose we wish to consider more complicated perturbation time dependences and remove the restriction that $g(x, t; \mu, \varepsilon)$ takes on the form in (2.2). We can still always introduce the associated autonomous system:

$$\begin{aligned}\dot{x} &= JDH(x) + \varepsilon g(x, \theta; \mu, \varepsilon) \\ \dot{\theta} &= 1.\end{aligned}\tag{6.1}$$

The phase space is now $X \times \mathbb{R}$ (the θ phase space is no longer compact), and we write $\theta = t + \theta_0$. An unperturbed homoclinic orbit in the autonomous system phase space is now as shown in figure 20(a). If one can establish for one's perturbation, as done in section 2 for quasiperiodic perturbations, persistence of one or more invariant normally hyperbolic sets, and the existence of infinitely many intersections of global stable and unstable manifolds in a given time slice $\chi(\bar{\theta}) = \{(x, \theta) \mid \theta = \bar{\theta}\}$ (and hence trivially in all time slices), then one has an invariant lobe structure upon which to define transport (see figure 20(b)). One can define an invariant core boundary and label the lobes in $X \times \mathbb{R}$ just as done in the quasiperiodic case (things will thus be piecewise continuous in θ). A sequence of monotonically increasing real numbers defines a sequence of sample times, and the evolution under (6.1) from one time slice to the next defines a sequence of maps. For any given time slice a two-dimensional lobe structure, a core boundary, and lobe labelling is defined by the intersection of the time slice with the invariant objects in $X \times \mathbb{R}$. Using the invariance of the lobe structure in the autonomous system phase space, one can monitor how the lobes map from one time slice to the next, identify turnstile lobes, and from this define and quantify entrainment and detrainment with respect to the time-dependent core, and so forth. One can numerically compute the lobe structure in any time slice using a method just like the *double phase slice method* (more

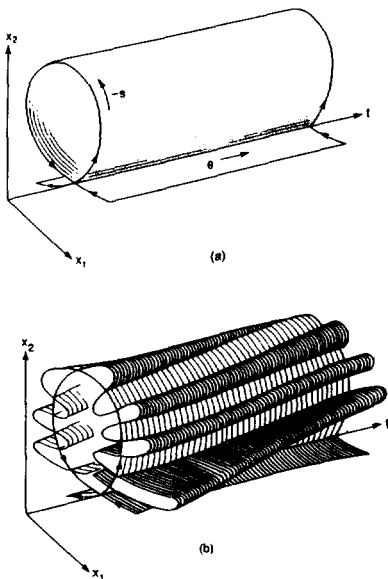


Figure 20. (a) An unperturbed homoclinic manifold in $X \times \mathbb{R}$; (b) an invariant lobe structure in $X \times \mathbb{R}$ for the perturbed case.

appropriately called a *double time slice method* here), which allows exact computation of flux and other transport quantities.

Meyer and Sell (1989) establish persistence of invariant normally hyperbolic manifolds for perturbations that are bounded and uniformly continuous in time, and then construct a Melnikov functional. Stoffer (1988a, b) defines a Melnikov function for non-autonomous maps and shows that the map sequence admits a hyperbolic set if the Melnikov function has infinitely many simple zeros with derivatives bounded away from zero. Their Melnikov analysis applying to (6.1) can be thought of as providing the following function:

$$M(s, \theta; \mu) = \int_{-\infty}^{\infty} DH(x_h(t)) \cdot g(x_h(t), t + s + \theta; \mu, \varepsilon = 0) dt \quad (6.2)$$

(see figure 20(a) for s and θ). The same arguments that lead to the expression for average flux (4.5) with the quasiperiodic Melnikov function (2.6) gives expression (4.5) with the Melnikov function (6.2) substituted:

$$\begin{aligned} \Phi &= \lim_{T \rightarrow \infty} \frac{\varepsilon}{2T} \int_0^T |M(s, \theta; \mu)| ds + O(\varepsilon^2) \\ &= \lim_{T \rightarrow \infty} \frac{\varepsilon}{2T} \int_0^T |M(s, \theta; \mu)| d\theta + O(\varepsilon^2). \end{aligned} \quad (6.3)$$

Thus the expression for average flux extends trivially. Note that, from (6.2), the first and second expressions of (6.3) will be independent of θ and s , respectively. In addition, the geometrical constraints of the two-dimensional lobes in any time slice are identical to those of the time-periodic case, so the transport equations of section 4 and Beigie *et al* (1991b) are identical under more general time dependences. As we see, then, the ideas for quasiperiodic perturbations are robust and extend easily to more general time dependences. For the more general perturbations, however, the global picture of the dynamics may be less apparent than in the quasiperiodic case, especially the two frequency case, since the θ variable is no longer compact. For instance, to find the turnstile lobes in each time slice one will truly have to deal with an infinite domain $\theta \in \mathbb{R}$, and one will not be able to give a global portrayal of the choice of τ_c and the invariant core boundary. Thus, because of their simplicity and geometric appeal, we wish to stress two frequency systems as workable and illuminating examples of chaotic transport under a bi-infinite sequence of non-repeating maps.

Acknowledgments

This material is based upon work supported by a National Science Foundation Graduate Fellowship, the National Science Foundation Presidential Young Investigator Program, the Office of Naval Research Young Investigator Program, and Caltech's Program in Advanced Technologies, sponsored by Aerojet General, General Motors, and TRW. We thank Peggy Firth and Cecilia Lin for the artwork.

Appendix. The quasiperiodically oscillating vortex pair (ovp) flow

Consider the quasiperiodic generalization of the oscillating vortex pair flow studied by Rom-Kedar *et al* (1990). This two-dimensional fluid flow consists of a pair of

point vortices of equal and opposite strength $\pm\Gamma$ in the presence of an oscillating strain-rate field. The streamfunction under the quasiperiodic forcing is, in the comoving frame,

$$\psi(x_1, x_2, t) = \frac{-\Gamma}{4\pi} \log \left(\frac{(x_1 - x_1^v(t))^2 + (x_2 - x_2^v(t))^2}{(x_1 - x_1^v(t))^2 + (x_2 + x_2^v(t))^2} \right) - V_v x_2 + \varepsilon x_1 x_2 \left(\sum_{i=1}^l \omega_i f_i \sin(\omega_i t + \theta_{i_0}) \right)$$

where $(x_1^v(t), \pm x_2^v(t))$ are the vortex positions, V_v is the average velocity of the vortex pair in the lab frame, and $\varepsilon \omega_i f_i$ is the strain rate amplitude associated with the i th frequency (εf_i is non-dimensional). For $\varepsilon = 0$, $(x_1^v, x_2^v) = (0, d)$ and $V_v = \Gamma/4\pi d$. The equations of motion for the vortices and passive marker particles are easily derived from this Hamiltonian (see Beigie *et al* (1991b) where the equations are expressed in *non-dimensional* form).

The equations of motion for x_1 and x_2 define a two-dimensional non-autonomous dynamical system, and the motion of (x_1, x_2) for a given Γ , ε , f_1 , f_2 , ω_1 , ω_2 , θ_{1_0} , θ_{2_0} and choice of initial vortex conditions is the fluid flow whose transport we address. The net perturbation is a sum of the forcing term, linear in ε , and the vortex response, nonlinear in ε ; by Taylor expanding the vortex term about $\varepsilon = 0$, the governing equations in the two frequency case can be put in the autonomous form

$$\begin{aligned} \dot{x} &= JDH(x) + \varepsilon g(x, \theta_1, \theta_2; f_1, f_2, \omega_1, \omega_2) + O(\varepsilon^2) \\ \dot{\theta}_1 &= \omega_1, \quad \dot{\theta}_2 = \omega_2 \end{aligned} \quad (A1)$$

(see Beigie *et al* 1991b). Figure 21(a) shows some level sets in $\Sigma^{\theta_{2_0}}$ of the unperturbed ($\varepsilon = 0$) system (A1). The two frequency Melnikov function is given by

$$\begin{aligned} M(s, \theta_1, \theta_{2_0}; f_1, f_2, \omega_1, \omega_2) \\ = f_1 \omega_1 F(\omega_1^{-1}) \sin(\omega_1 s + \theta_1) + f_2 \omega_2 F(\omega_2^{-1}) \sin(\omega_2 s + \theta_{2_0}) \end{aligned} \quad (A2)$$

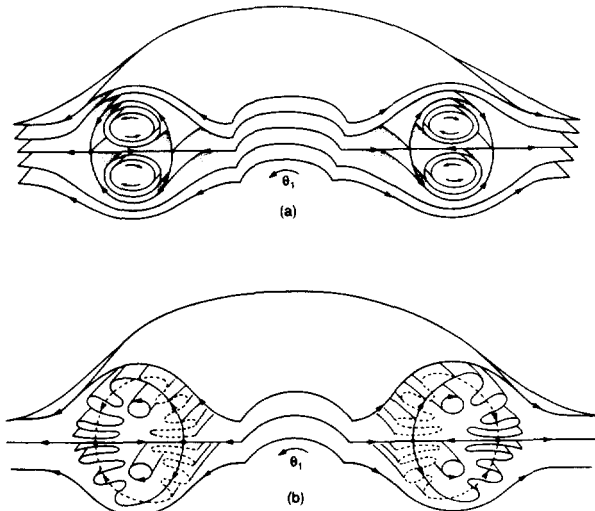


Figure 21. (a) Some level sets in $\Sigma^{\theta_{2_0}}$ of the unperturbed OVP flow. (b) The lobe structure in $\Sigma^{\theta_{2_0}}$ of the perturbed OVP flow.

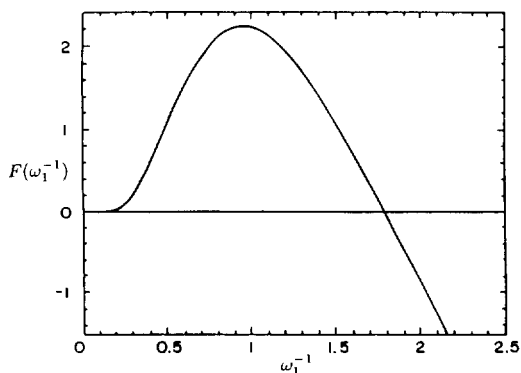


Figure 22. The plot of $F(\omega_1^{-1})$ (taken from Rom-Kedar *et al* (1990)).

(see Beigie *et al* 1991b), where $F(\omega^{-1})$ is shown in figure 22. From section 2 we refer to the ratio of each Melnikov function amplitude $f_i\omega_i F(\omega_i^{-1})$ to the corresponding relative perturbation amplitude $\omega_i f_i$ as the *relative scaling factor* associated with frequency ω_i . The fact that the relative scaling factors $F(\omega_i^{-1})$ are frequency dependent is pertinent to the study of transport rates, in particular to a comparison of average flux between single and multiple frequency forcing. Note how all scaling factors are determined by the single *relative scaling function* $F(\omega^{-1})$. The Melnikov function is easily seen to have zero sets in all phase slices of $\Sigma^{\theta_{20}}$, and hence the upper and lower stable and unstable manifolds in $\Sigma^{\theta_{20}}$ of the perturbed flow form the boundary of the defined three-dimensional heteroclinic lobe structure (see figure 21(b)) and intersect each phase slice to give the boundary of a two-dimensional lobe structure. The physical situation can be pictured as the two-dimensional fluid mapping from one time slice to the next in $\Sigma^{\theta_{20}}$ with each successive time sample. The lobe structure in $\Sigma^{\theta_{20}}$ thus constrains the motion of the fluid in the tangle region, as described in section 3.

References

- Beigie D, Leonard A and Wiggins S 1991a A global study of enhanced stretching and diffusion in chaotic tangles *Phys. Fluids A* **3** 1039–50
- 1991b The dynamics associated with the chaotic tangles of two dimensional quasiperiodic vector fields: theory and applications *Nonlinear Phenomena in Atmospheric and Oceanic Sciences* ed G F Carnevale and R Pierrehumbert (Berlin: Springer) in press
- Beigie D and Wiggins S 1991 The dynamics associated with a quasiperiodically forced Morse oscillator: application to molecular dissociation *Caltech preprint*
- Bensimon D and Kadanoff L P 1984 Extended chaos and disappearance of KAM trajectories *Physica* **13D** 82–9
- Fenichel N 1971 Persistence and Smoothness of Invariant Manifolds for Flows *Ind. Univ. Math. J.* **21** 193–226
- Gillilan R E and Ezra G S 1991 Transport and turnstiles in multidimensional Hamiltonian mappings for unimolecular fragmentation: Application to van der Waals predissociation *J. Chem. Phys.* **94** 2648–68
- Goggin M E and Milonni P W 1988a Driven Morse oscillator: classical chaos, quantum theory, and photodissociation *Phys. Rev. A* **37** 796–806
- 1988b Driven Morse oscillator: classical chaos and quantum theory for two-frequency excitation *Phys. Rev. A* **38** 5174–81
- Guckenheimer J and Holmes P 1983 *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields* (Berlin: Springer)

- Hirsch M W, Pugh C C and Shub M 1977 *Invariant Manifolds (Springer Lecture Notes in Mathematics 583)* (Berlin: Springer)
- Ide K and Wiggins S 1989 The bifurcation to homoclinic tori in the quasiperiodically forced Duffing oscillator *Physica* **34D** 169–82
- MacKay R S and Meiss J D 1988 Relation between quantum and classical thresholds for multiphoton ionization of excited atoms *Phys. Rev. A* **37** 4702–6
- MacKay R S, Meiss J D and Percival I C 1984 Transport in Hamiltonian systems *Physica* **13D** 55–81
- Meyer K R and Sell G R 1989 Melnikov transforms, Bernoulli bundles, and almost periodic perturbations *Trans. Am. Math. Soc.* **314** 63–105
- Moon F C and Holmes W T 1985 Double Poincaré sections of a quasi-periodically forced, chaotic attractor *Phys. Lett* **111A** 157–60
- Noid D W and Stine J R 1979 Infrared multiphoton dissociation with one and two lasers *Chem. Phys. Lett.* **65** 153–7
- Rom-Kedar V, Leonard A and Wiggins S 1990 An analytical study of transport, mixing and chaos in an unsteady vortical flow *J. Fluid Mech.* **214** 347–94
- Rom-Kedar V and Wiggins S 1990 Transport in Two-Dimensional Maps *Arch. Rat. Mech. Anal.* **109** 3, 239–98
- Scheurle J 1986 Chaotic solutions of systems with almost periodic forcing *J Appl. Math. Phys. (ZAMP)* **37** 12–26
- Stoffer D 1988a Transversal homoclinic points and hyperbolic sets for non-autonomous maps I. *J. Appl. Math. Phys. (ZAMP)* **39** 518–49
- 1988b Transversal homoclinic points and hyperbolic sets for non-autonomous maps II *J. Appl. Math. Phys. (ZAMP)* **39** 783–812
- Wiggins S 1987 Chaos in the quasiperiodically forced Duffing oscillator *Phys. Lett.* **124A** 138–42
- 1988 *Global Bifurcations and Chaos—Analytical Methods* (Berlin: Springer)
- 1990 On the Geometry of Transport in Phase Space, I. Transport in k -Degree-of-Freedom Hamiltonian Systems, $2 \leq k < \infty$ *Physica* **44D** 471–501